

# Bivariate least squares linear regression: towards a unified analytic formalism.

## I. Functional models

R. Caimmi\*

January 13, 2013

### Abstract

Concerning bivariate least squares linear regression, the classical approach pursued for functional models in earlier attempts (York, 1966; 1969) is reviewed using a new formalism in terms of deviation (matrix) traces which, for homoscedastic data, reduce to usual quantities leaving aside an unessential (but dimensional) multiplicative factor. Within the framework of classical error models, the dependent variable relates to the independent variable according to the usual additive model. The classes of linear models considered are regression lines in the general case of correlated errors in  $X$  and in  $Y$  for heteroscedastic data, and in the opposite limiting situations of (i) uncorrelated errors in  $X$  and in  $Y$ , and (ii) completely correlated errors in  $X$  and in  $Y$ . The special case of (C) generalized orthogonal regression is considered in detail together with well known subcases, namely: (Y) errors in  $X$  negligible (ideally null) with respect to errors in  $Y$ ; (X) errors in  $Y$  negligible (ideally null) with respect to errors in  $X$ ; (O) genuine orthogonal regression; (R) reduced major-axis regression. In the limit of homoscedastic data, the results determined for functional models are compared with their counterparts related to extreme structural models i.e. the instrumental scatter is negligible (ideally null) with respect

---

\* *Astronomy Department, Padua Univ., Vicolo Osservatorio 3/2, I-35122 Padova, Italy*  
email: roberto.caimmi@unipd.it fax: 39-049-8278212

to the intrinsic scatter (Isobe et al., 1990; Feigelson and Babu, 1992). While regression line slope and intercept estimators for functional and structural models necessarily coincide, the contrary holds for related variance estimators even if the residuals obey a Gaussian distribution, with the exception of Y models. An example of astronomical application is considered, concerning the  $[O/H]$ - $[Fe/H]$  empirical relations deduced from five samples related to different stars and/or different methods of oxygen abundance determination. For selected samples and assigned methods, different regression models yield consistent results within the errors ( $\mp\sigma$ ) for both heteroscedastic and homoscedastic data. Conversely, samples related to different methods produce discrepant results, due to the presence of (still undetected) systematic errors, which implies no definitive statement can be made at present. A comparison is also made between different expressions of regression line slope and intercept variance estimators, where fractional discrepancies are found to be not exceeding a few percent, which grows up to about 20% in presence of large dispersion data. An extension of the formalism to structural models is left to a forthcoming paper.

*keywords - galaxies: evolution - stars: formation; evolution - methods: data analysis - methods: statistical.*

pacs codes: 98.62.-g; 97.10.Cv; 02.50.-r

## 1 Introduction

Linear regression is a fundamental and frequently used statistical tool in almost all branches of science, such as astronomy, biology, chemistry, geology, physics, and statistics of course; for a full discussion refer to a classical paper (Isobe et al., 1990; hereafter quoted as Ial90). In spite of its apparent simplicity, the task of drawing the “best” straight line through data on a Cartesian plot is difficult and controversial. The problem is twofold: regression line slope and intercept estimators are expressed involving minimizing or maximizing some function of the data; on the other hand, regression line slope and intercept variance estimators are expressed requiring knowledge of the error distributions of the data.

The complexity mainly arises from the occurrence of intrinsic dispersion, which could be related to a non Gaussian distribution, in addition to the dispersion related to the measurement processes (hereafter quoted as instrumental dispersion), which necessarily implies a Gaussian distribution. An increasing difficulty is encountered in more exotic situations, such as truncated regression, where a variable is assumed to be truncated below or above a threshold, and censored regression, where several data are assumed to be

undetected at various sensitivity levels. For further details of astronomical interest, refer to a classical paper (Feigelson and Babu, 1992; erratum, 2011; hereafter quoted together as FB92) and, in general, to specific texts on the subject (e.g., Klein and Moeschberger, 2005).

In statistics, problems where the true points lie precisely on an expected line are called functional regression models, while problems where the true points are (intrinsically) scattered about an expected line are called structural regression models. Accordingly, functional regression models may be conceived as structural regression models where the intrinsic dispersion is negligible (ideally null) with respect to the instrumental dispersion. A distinction between functional and structural modelling is currently preferred, where the former can be affected by intrinsic scatter but with no or only minimal assumptions on related distributions, while the latter implies (usually parametric) models are placed on the above mentioned distributions. For further details refer to specific textbooks (e.g., Carroll et al., 2006, Chap. 2, §2.1). In addition, models where the instrumental dispersion is the same from point to point for each variable, are called homoscedastic models, while models where the instrumental dispersion is (in general) different from point to point, are called heteroscedastic models. Similarly, related data are denoted as homoscedastic and heteroscedastic, respectively.

Bivariate least squares linear regression related to heteroscedastic functional models with uncorrelated and correlated errors, following Gaussian distributions, were analysed and formulated in two classical papers (York, 1966; 1969; hereafter quoted as Y66 and Y69, respectively). Bivariate least squares linear regression related to extreme structural models, where the instrumental dispersion is negligible (ideally null) with respect to intrinsic dispersion, was exhaustively treated in a classical paper (Ial90).

An extension to homoscedastic functional and structural models was performed in a subsequent paper (FB92), yielding the same expression of regression line slope and intercept estimators, provided the instrumental dispersion in the former case coincides with the intrinsic dispersion (assumed to be dominant) in the latter case, for each variable, and the residuals follow a Gaussian distribution.

Further extension to homoscedastic structural models where instrumental and intrinsic dispersion are of the same order, was carried in a later paper (Akritas and Bershday, 1996; hereafter quoted as AB96). Heteroscedastic structural models with instrumental dispersion negligible with respect to intrinsic dispersion in one variable, were also presented (AB96).

The above mentioned papers provide the simplest description of linear regression. More sophisticated attempts imply additional effects such as truncated and censored regression (e.g., FB92), analytical methods such as

correction of the observed moments of the data (e.g., Fuller, 1987; AB96; Freedman et al., 2004), minimization of an effective  $\chi^2$  statistic (e.g., Clutton-Brock, 1967; Barker and Diana, 1974; Press et al., 1992; Tremaine et al., 2002), assuming a probability distribution for the true independent variable values (e.g., Schafer, 1987, 2001; Roy and Banerjee, 2006), computational methods such as bootstrap and jackknife (e.g., FB92), matrix formalism (e.g., Schwarzenberg-Czerny, 1995; Branham, 2001), simultaneous adjustment (e.g., Pourbaix, 1998), and Bayesian approach (e.g., Zellner, 1971; Gull, 1989; Dellaportas and Stephens, 1995; Carroll et al., 1999; Scheines et al., 1999; Kelly, 2007).

The last investigation is particularly relevant in that it is the first example, in the astronomical literature, where linear regression is considered following the modern (since about half a century ago) approach based on likelihoods rather than the old (up to about a century ago) least-squares approach. More specifically, a hierarchical measurement error model is set up therein, the complicated likelihood is written down, and a variety of minimum least-squares and Bayesian solutions are shown, which can treat functional, structural, multivariate, truncated and censored measurement error regression problems.

Even in dealing with the simplest homoscedastic (or heteroscedastic) functional and structural models, still no unified analytic formalism has been developed (to the knowledge of the author) where (i) structural heteroscedastic models with instrumental and intrinsic dispersion of comparable order in both variables, are considered; (ii) previous results are recovered in the limit of dominant instrumental dispersion; and (iii) previous results are recovered in the limit of dominant intrinsic dispersion. A related formulation may be useful also for computational methods, in the sense that both the general case and limiting situations can be described by a single numerical code.

The current paper aims at making a first step towards a unified analytic formalism of bivariate least squares linear regression involving functional models. More specifically, earlier attempts shall be reviewed and reformulated by definition and use of deviation (matrix) traces, within the framework of classical error models where the dependent variable relates to the independent variable according to the usual additive model.

Homoscedastic and heteroscedastic functional models are presented in section 2, basing on two classical papers (Y66; Y69). An example of astronomical application is outlined in section 3. The discussion is performed in section 4. Finally, the conclusion is shown in section 5. Some points are developed with more detail in the Appendix. An extension of the formalism to structural models is left to a forthcoming paper.

## 2 Least-squares fitting of a straight line

### 2.1 General considerations

Attention shall be restricted to the classical problem of least-squares fitting of a straight line, where both variables are measured with errors and the true points lie on the unknown regression line i.e. functional models (e.g., Y66; Y69), which can be considered as structural models in the limit of negligible (ideally null) intrinsic scatter. In general, the dependent variable,  $y$ , relates to the independent variable,  $x$ , according to the usual additive model (e.g., AB96; Carroll et al., 2006, Chap. 1, §1.2, Chap. 3, §3.2.1; Kelly, 2007; Buonaccorsi, 2010, Chap. 4, §4.3):

$$y_{Si} = ax_{Si} + b + \epsilon_i \quad ; \quad 1 \leq i \leq n \quad ; \quad (1)$$

where  $P_{Si}^* \equiv (x_{Si}, y_{Si})$  are the actual points whose coordinates are affected by no instrumental error and  $\epsilon_i$  is a random variable with null expectation value representing the intrinsic scatter in  $(x_{Si}, y_{Si})$  about the regression line<sup>1</sup>.

Due to the occurrence of instrumental errors, the observed points,  $P_i \equiv (X_i, Y_i)$ , are evaluated in place of the actual points,  $P_{Si}^*$ . The coordinates of observed and actual points are assumed to be related as:

$$X_i = x_{Si} + (\xi_{Fx})_i \quad ; \quad 1 \leq i \leq n \quad ; \quad (2a)$$

$$Y_i = y_{Si} + (\xi_{Fy})_i \quad ; \quad 1 \leq i \leq n \quad ; \quad (2b)$$

where  $(\xi_{Fx})_i$ ,  $(\xi_{Fy})_i$ , are the instrumental errors on  $x_{Si}$  and  $y_{Si}$ , respectively, assumed to be normally distributed with null expectation values and known variances,  $(\sigma_{xx})_i = [(\sigma_x)_i]^2$ ,  $(\sigma_{yy})_i = [(\sigma_y)_i]^2$ , and covariance,  $(\sigma_{xy})_i = (\sigma_{yx})_i$ . The terms “independent variable” and “dependent variable” are purely conventional when the model is symmetric in  $x$  and in  $y$  provided  $a \neq 0$ . For a vanishing intrinsic scatter,  $\epsilon_i \rightarrow 0$ ,  $1 \leq i \leq n$ , actual points lie on the unknown regression line whereas adjusted points,  $\hat{P}_i \equiv (x_i, y_i)$ , lie on the estimated regression line:

$$y_i = \hat{a}x_i + \hat{b} \quad ; \quad 1 \leq i \leq n \quad ; \quad (3)$$

where, in general, estimators are denoted by hats. For further details refer to earlier attempts (York, 1967; Y69).

In the case under discussion, the regression estimator minimizes the sum (over the  $n$  observations) of squared residuals (e.g., Y69), or statistical distances of the observed points,  $P_i \equiv (X_i, Y_i)$ , from the estimated line in the unknown parameters,  $a, b, x_1, \dots, x_n$  (e.g., Fuller, 1987, Chap. 1, §1.3.3). Under

---

<sup>1</sup>The Italian convention shall be adopted here, according to which the slope and the intercept of a straight line on the Cartesian plane, are denoted as  $a$ ,  $b$ , respectively.

restrictive assumptions, the regression estimator is the functional maximum likelihood estimator (e.g., Carroll et al., 2006, Chap. 3, §3.4.2).

To the knowledge of the author, only classical error models are considered for astronomical applications, and for this reason different error models such as Berkson models and mixture error models (e.g., Carroll et al., 2006, Chap. 3, Sect. 3.2) shall not be dealt with in the current attempt. From this point on, investigation shall be limited to functional models and least-squares regression estimators for the following reasons. First, they are important models in their own right, furnishing an approximation to real world situations. Second, a careful examination of these simple models helps for understanding the theoretical underpinnings of methods for other models of greater complexity such as hierarchical models (e.g., Kelly, 2007).

## 2.2 Functional models

With regard to functional models, bivariate least squares linear regression were analysed in two classical papers (Y66; Y69). The same line of thought shall be followed here and the sole changes shall be concerned with the formalism, as clearly indicated. The general case shall first be presented, while special cases shall be deduced later as limiting situations.

In the light of the model outlined in subsection 2.1 in absence of intrinsic scatter, Eqs. (2) and (3), the actual points,  $P_{Si}^* \equiv (x_{Si}, y_{Si})$  coincide with the true points,  $P_i^* \equiv (x_i^*, y_i^*)$ , whose coordinates lie on the unknown regression line:

$$y_i^* = ax_i^* + b \quad ; \quad 1 \leq i \leq n \quad ; \quad (4)$$

while the observed points,  $P_i \equiv (X_i, Y_i)$ , are scattered with respect to the regression line.

The coordinates,  $(x_i, y_i)$ , may be conceived as the adjusted values of related observations,  $(X_i, Y_i)$ , on the calculated regression line (Y66; Y69), Eq. (3) and, in addition, as estimators of the coordinates,  $(x_i^*, y_i^*)$ , on the true regression line determined in absence of measurement errors, Eq. (4). The line of adjustment,  $\overline{P_i \hat{P}_i}$  (e.g., Y69), may be conceived as an estimator of the statistical distance,  $\overline{P_i P_i^*}$  (e.g., Fuller, 1987, Chap. 1, §1.3.3), where  $\hat{P}_i(x_i, y_i)$  is the adjusted point on the estimated regression line and  $P_i^*(x_i^*, y_i^*)$  is the true point on the true regression line.

The squared weighted residuals are defined as (Y69):

$$(\tilde{R}_i)^2 = \frac{w_{x_i}(X_i - x_i)^2 + w_{y_i}(Y_i - y_i)^2 - 2r_i\sqrt{w_{x_i}w_{y_i}}(X_i - x_i)(Y_i - y_i)}{1 - r_i^2} \quad ; \quad (5a)$$

$$r_i = \frac{(\sigma_{xy})_i}{[(\sigma_{xx})_i(\sigma_{yy})_i]^{1/2}} \quad ; \quad |r_i| \leq 1 \quad ; \quad 1 \leq i \leq n \quad ; \quad (5b)$$

where  $w_{x_i}, w_{y_i}$ , are the weights of the various measurements (or observations) and  $r_i$  the correlation coefficients. An equivalent formulation in matrix formalism can be found in specific textbooks, where weighted true residuals are conceived as “statistical distances” from data points to related points on the regression line [e.g., Fuller, 1987, Chap. 1, §1.3.3, Eq. (1.3.16)].

In the limit of uncorrelated errors, Eq. (5) reduces to (Y66):

$$(\tilde{R}_i)^2 = w_{x_i}(X_i - x_i)^2 + w_{y_i}(Y_i - y_i)^2 ; \quad (6a)$$

$$r_i = 0 ; \quad 1 \leq i \leq n ; \quad (6b)$$

where the covariances are necessarily null,  $(\sigma_{xy})_i = 0, 1 \leq i \leq n$ .

In the limit of perfectly correlated errors,  $r_i = \text{sgn}(r_i), 1 \leq i \leq n$ , it can be seen that the following relation holds:

$$\frac{Y_i - y_i}{X_i - x_i} = \left( \frac{w_{x_i}}{w_{y_i}} \right)^{1/2} \text{sgn}(r_i) ; \quad (7)$$

where  $\text{sgn}$  is the sign function<sup>2</sup>. Accordingly, Eq. (5a) reduces to:

$$\begin{aligned} (\tilde{R}_i)^2 &= \lim_{r_i \rightarrow \text{sgn}(r_i)} \left[ \frac{w_{x_i}(X_i - x_i)^2 + w_{x_i}(X_i - x_i)^2 \text{sgn}^2(r_i)}{1 - r_i^2} \right. \\ &\quad \left. - \frac{2r_i \text{sgn}(r_i) w_{x_i}(X_i - x_i)^2}{1 - r_i^2} \right] \\ &= w_{x_i}(X_i - x_i)^2 \lim_{r_i \rightarrow \text{sgn}(r_i)} \frac{2 - 2r_i \text{sgn}(r_i)}{1 - r_i^2 \text{sgn}^2(r_i)} \\ &= w_{x_i}(X_i - x_i)^2 \lim_{r_i \rightarrow \text{sgn}(r_i)} \frac{2}{1 + r_i \text{sgn}(r_i)} ; \end{aligned}$$

which, owing to Eq. (7), takes the form:

$$(\tilde{R}_i)^2 = w_{x_i}(X_i - x_i)^2 = w_{y_i}(Y_i - y_i)^2 ; \quad (8a)$$

$$r_i = \text{sgn}(r_i) ; \quad 1 \leq i \leq n ; \quad (8b)$$

where the covariances are necessarily equal to the (positive or negative) square root of the variance product,  $(\sigma_{xy})_i = [(\sigma_{xx})_i(\sigma_{yy})_i]^{1/2}, 1 \leq i \leq n$ .

Turning back to the general case, let the squared residual matrix be defined as:

$$M_{\tilde{R}} = || \tilde{R}_i \tilde{R}_j || ; \quad 1 \leq i \leq n ; \quad 1 \leq j \leq n ; \quad (9)$$

---

<sup>2</sup>The sign function is defined as  $\text{sgn}(x) = |x|/x, x \neq 0; \text{sgn}(x) = 0, x = 0$ .

which is a square matrix of order,  $n$ . Let the related trace:

$$T_{\tilde{R}} = \sum_{i=1}^n (\tilde{R}_i)^2 ; \quad (10)$$

be defined as the squared residual trace. The regression estimator is that minimizing the squared residual trace (Y66; Y69):

$$\begin{aligned} T_{\tilde{R}}(x_1, x_2, \dots, x_n, y_1, y_2, \dots, y_n) &= \sum_{i=1}^n (\tilde{R}_i)^2 \\ &= \sum_{i=1}^n \frac{w_{x_i}(X_i - x_i)^2 + w_{y_i}(Y_i - y_i)^2 - 2r_i\sqrt{w_{x_i}w_{y_i}}(X_i - x_i)(Y_i - y_i)}{1 - r_i^2} ; \end{aligned} \quad (11)$$

with the constraint expressed by Eq. (3) where the coefficients,  $\hat{a}$ ,  $\hat{b}$ , are still to be determined, and for this reason are denoted as  $a$ ,  $b$ , respectively. If the values,  $(x_1, x_2, \dots, x_n, y_1, y_2, \dots, y_n, a, b)$ , relate to a constrained extremum point, then the following relations must necessarily hold (Y66; Y69):

$$\delta T_{\tilde{R}} = -2 \sum_{i=1}^n \left[ \frac{w_{x_i}(X_i - x_i)\delta x_i + w_{y_i}(Y_i - y_i)\delta y_i}{1 - r_i^2} - \frac{r_i\sqrt{w_{x_i}w_{y_i}}[(Y_i - y_i)\delta x_i + (X_i - x_i)\delta y_i]}{1 - r_i^2} \right] = 0 ; \quad (12)$$

$$\delta y_i - a\delta x_i - x_i\delta a - \delta b = 0 ; \quad 1 \leq i \leq n ; \quad (13)$$

where Eq. (13) is also valid after inserting on both sides a multiplier (to be specified later),  $\lambda_i$ ,  $1 \leq i \leq n$ . The sum of the ensuing relations yields:

$$\sum_{i=1}^n \lambda_i \delta y_i - a \sum_{i=1}^n \lambda_i \delta x_i - \sum_{i=1}^n \lambda_i x_i \delta a - \sum_{i=1}^n \lambda_i \delta b = 0 ; \quad (14)$$

and the sum of the left-hand sides of Eqs. (12) and (14) produces:

$$\begin{aligned} &\sum_{i=1}^n \left[ \frac{w_{x_i}(X_i - x_i) - r_i\sqrt{w_{x_i}w_{y_i}}(Y_i - y_i)}{1 - r_i^2} - a\lambda_i \right] \delta x_i - \sum_{i=1}^n \lambda_i \delta b \\ &+ \sum_{i=1}^n \left[ \frac{w_{y_i}(Y_i - y_i) - r_i\sqrt{w_{x_i}w_{y_i}}(X_i - x_i)}{1 - r_i^2} + \lambda_i \right] \delta y_i - \sum_{i=1}^n \lambda_i x_i \delta a = 0 ; \end{aligned} \quad (15)$$

which implies each coefficient is null, as:

$$\frac{w_{x_i}}{1 - r_i^2}(X_i - x_i) - \frac{r_i\sqrt{w_{x_i}w_{y_i}}}{1 - r_i^2}(Y_i - y_i) - a\lambda_i = 0 ; \quad (16a)$$

$$\frac{w_{y_i}}{1 - r_i^2}(Y_i - y_i) - \frac{r_i\sqrt{w_{x_i}w_{y_i}}}{1 - r_i^2}(X_i - x_i) + \lambda_i = 0 ; \quad (16b)$$



$$\sum_{i=1}^n \lambda_i = 0 \quad ; \quad (17)$$

$$\sum_{i=1}^n \lambda_i x_i = 0 \quad ; \quad (18)$$

for  $1 \leq i \leq n$ .

The combination of Eqs. (16a) and (16b) yields:

$$X_i - x_i = \left( a - \frac{r_i}{\Omega_i} \right) \frac{\lambda_i}{w_{x_i}} \quad ; \quad (19a)$$

$$Y_i - y_i = \frac{r_i}{\Omega_i} \left( a - \frac{1}{r_i \Omega_i} \right) \frac{\lambda_i}{w_{x_i}} \quad ; \quad (19b)$$

$$\Omega_i = \sqrt{\frac{w_{y_i}}{w_{x_i}}} \quad ; \quad 1 \leq i \leq n \quad ; \quad (20)$$

and the combination of Eqs. (3) and (19) produces:

$$\lambda_i = W_i(aX_i + b - Y_i) \quad ; \quad 1 \leq i \leq n \quad ; \quad (21)$$

$$W_i = \frac{w_{x_i} \Omega_i^2}{1 + a^2 \Omega_i^2 - 2a r_i \Omega_i} \quad ; \quad 1 \leq i \leq n \quad ; \quad (22)$$

finally, the substitution of Eq. (21) into (17) and (18) yields, in the latter case after some algebra:

$$\sum_{i=1}^n W_i(aX_i + b - Y_i) = 0 \quad ; \quad (23)$$

$$\begin{aligned} & \sum_{i=1}^n W_i X_i (aX_i + b - Y_i) - a \sum_{i=1}^n \frac{W_i^2}{w_{x_i}} (aX_i + b - Y_i)^2 \\ & + \sum_{i=1}^n \frac{W_i^2}{w_{x_i}} \frac{r_i}{\Omega_i} (aX_i + b - Y_i)^2 = 0 \quad ; \end{aligned} \quad (24)$$

where the regression line slope and intercept estimators,  $\hat{a}$  and  $\hat{b}$ , are found solving the system of Eqs. (23) and (24).

In terms of the weighted means:

$$\tilde{Z} = \frac{\sum_{i=1}^n W_i Z_i}{\sum_{i=1}^n W_i} \quad ; \quad Z = X, Y \quad ; \quad (25a)$$

$$\sum_{i=1}^n W_i (Z_i - \tilde{Z}) = 0 \quad ; \quad Z = X, Y \quad ; \quad (25b)$$

the intercept is expressed by casting Eq. (23) under the equivalent form:

$$b = \tilde{Y} - a\tilde{X} \quad ; \quad (26)$$

the point,  $\tilde{\mathbf{P}} \equiv (\tilde{X}, \tilde{Y})$ , clearly lies on the regression line, and can be considered as the “barycentre” of the data,  $\mathbf{P}_i \equiv (X_i, Y_i)$ ,  $1 \leq i \leq n$  (Y66; Y69).

Using Eq. (26), the following relation holds:

$$aX_i + b - Y_i = a(X_i - \tilde{X}) - (Y_i - \tilde{Y}) \quad ; \quad 1 \leq i \leq n \quad ; \quad (27)$$

in terms of the deviations from the weighted mean,  $(X_i - \tilde{X})$  and  $(Y_i - \tilde{Y})$ . Using Eqs. (23), (26) and (27), together with the identities,  $Z_i = (Z_i - \tilde{Z}) + \tilde{Z}$ ,  $1 \leq i \leq n$ ,  $Z = X, Y$ , the following relations are found after some algebra:

$$\sum_{i=1}^n W_i X_i (aX_i + b - Y_i) = a \sum_{i=1}^n W_i (X_i - \tilde{X})^2 - \sum_{i=1}^n W_i (X_i - \tilde{X})(Y_i - \tilde{Y}); \quad (28)$$

$$\begin{aligned} \sum_{i=1}^n V_i (aX_i + b - Y_i)^2 &= a^2 \sum_{i=1}^n V_i (X_i - \tilde{X})^2 + \sum_{i=1}^n V_i (Y_i - \tilde{Y})^2 \\ &\quad - 2a \sum_{i=1}^n V_i (X_i - \tilde{X})(Y_i - \tilde{Y}) \quad ; \end{aligned} \quad (29)$$

$$\begin{aligned} \sum_{i=1}^n U_i (aX_i + b - Y_i)^2 &= a^2 \sum_{i=1}^n U_i (X_i - \tilde{X})^2 + \sum_{i=1}^n U_i (Y_i - \tilde{Y})^2 \\ &\quad - 2a \sum_{i=1}^n U_i (X_i - \tilde{X})(Y_i - \tilde{Y}) \quad ; \end{aligned} \quad (30)$$

$$V_i = \frac{W_i^2}{w_{x_i}} = \frac{w_{x_i} \Omega_i^4}{(1 + a^2 \Omega_i^2 - 2ar_i \Omega_i)^2} \quad ; \quad 1 \leq i \leq n \quad ; \quad (31)$$

$$U_i = \frac{W_i^2}{w_{x_i}} \frac{r_i}{\Omega_i} = \frac{w_{x_i} r_i \Omega_i^3}{(1 + a^2 \Omega_i^2 - 2ar_i \Omega_i)^2} \quad ; \quad 1 \leq i \leq n \quad ; \quad (32)$$

and the substitution of Eqs. (28), (29), (30), into (24) yields (Y69):

$$\begin{aligned} &a^3 \sum_{i=1}^n V_i (X_i - \tilde{X})^2 - a^2 \left[ 2 \sum_{i=1}^n V_i (X_i - \tilde{X})(Y_i - \tilde{Y}) + \sum_{i=1}^n U_i (X_i - \tilde{X})^2 \right] \\ &- a \left[ \sum_{i=1}^n W_i (X_i - \tilde{X})^2 - \sum_{i=1}^n V_i (Y_i - \tilde{Y})^2 - 2 \sum_{i=1}^n U_i (X_i - \tilde{X})(Y_i - \tilde{Y}) \right] \\ &+ \left[ \sum_{i=1}^n W_i (X_i - \tilde{X})(Y_i - \tilde{Y}) - \sum_{i=1}^n U_i (Y_i - \tilde{Y})^2 \right] = 0 \quad ; \end{aligned} \quad (33)$$

where the resulting terms have been ordered in decreasing powers of the slope,  $a$ , and the coefficients,  $W_i$ ,  $V_i$ ,  $U_i$ ,  $1 \leq i \leq n$ , depend in turn on the slope, as

shown by Eqs. (22), (31), (32), respectively. Then Eq. (33) is a pseudo cubic equation which can be iteratively solved with the desired degree of precision, provided  $W_i, V_i, U_i, 1 \leq i \leq n$ , are weakly dependent on  $a$ .

With the aim of getting a more compact formalism, let the weighted deviation matrices be defined as:

$$M_{\tilde{Q}_{pq}} = \left\| \sqrt{Q_i Q_j} (X_i - \tilde{X})^p (Y_j - \tilde{Y})^q \right\| ; \quad 1 \leq i \leq n ; \quad 1 \leq j \leq n ; \quad (34)$$

which are square matrices of order,  $n$ . Let the related traces:

$$\tilde{Q}_{pq} = \sum_{i=1}^n Q_i (X_i - \tilde{X})^p (Y_i - \tilde{Y})^q ; \quad Q = W, V, U, P ; \quad (35)$$

be defined as the weighted deviation traces. Pure and mixed traces occur for  $p = 0$  and/or  $q = 0$ ;  $p > 0$  and  $q > 0$ ; respectively. The special cases,  $(p, q) = (2, 0), (0, 2); (1, 1)$ ; relate to expressions used in earlier attempts for both weighted (Y66; Y69) and unweighted (Ial90; FB92) residuals. The special case,  $(p, q) = (0, 0)$ , yields the product,  $n\overline{Q}$ . With this notation, Eq. (33) reads:

$$\tilde{V}_{20}a^3 - (2\tilde{V}_{11} + \tilde{U}_{20})a^2 - (\tilde{W}_{20} - \tilde{V}_{02} - 2\tilde{U}_{11})a + (\tilde{W}_{11} - \tilde{U}_{02}) = 0 ; \quad (36)$$

and the solutions can be determined along the following steps.

- (1) Estimate the slope,  $a^{(1)}$ , of the regression line, and calculate related deviation traces,  $\tilde{Q}_{pq}^{(1)}$ , appearing in Eq. (36).
- (2) Solve the corresponding cubic equation, and select the solution of interest,  $a^{(2)}$ .
- (3) Check the inequality,  $|a^{(i)}/a^{(i-1)} - 1| < v$ , where  $i = 2$  and  $v$  is the desired degree of precision.
- (4) If the above inequality is not satisfied, return to (1) using  $a^{(2)}$  instead of  $a^{(1)}$ , or in general  $a^{(i+1)}$  instead of  $a^{(i)}$ .
- (5) If the above inequality is satisfied for  $i = n$ , then  $\hat{a} = a^{(n)}$  is the regression line slope estimator. As pointed out in the parent paper (Y66) and confirmed by the author, in all cases considered three real solutions are found, where the one of interest coincides with the third appearing in standard formulae. For further details refer to Appendix A.

The pseudo cubic, Eq. (36), may be reduced algebraically to a pseudo quadratic equation (Y69) which, in terms of deviation traces, expressed by Eq. (35) via Eqs. (31) and (32), may be cast under the form:

$$(\tilde{V}_{11} - \tilde{U}_{20})a^2 + (\tilde{P}_{20} - \tilde{V}_{02})a - (\tilde{P}_{11} - \tilde{U}_{02}) = 0 \quad ; \quad (37)$$

$$P_i = \frac{W_i^2}{w_{y_i}} = \frac{W_i^2}{w_{x_i}\Omega_i^2} = \frac{w_{x_i}\Omega_i^2}{(1 + a^2\Omega_i^2 - 2ar_i\Omega_i)^2} \quad ; \quad 1 \leq i \leq n \quad ; \quad (38)$$

where no explanation and no quotation are provided in the parent paper (Y69) on how Eq. (37) can be deduced from Eq. (36). Surely, the solution of interest must necessarily be chosen among the three of the pseudo cubic or the two of the pseudo quadratic.

At this stage, the regression line slope and intercept estimators,  $\hat{a}$ , and  $\hat{b}$ , can be determined via Eqs. (36) or (37) and (26), respectively. An exact expression of the regression line slope and intercept variance estimators may be calculated using the method of partial differentiation, along the following steps (Y69).

- (1) Cast the pseudo quadratic equation, Eq. (37), into the implicit form:

$$\phi(X_i, Y_i, \hat{a}) = A(X_i, Y_i, \hat{a})(\hat{a})^2 + B(X_i, Y_i, \hat{a})\hat{a} + C(X_i, Y_i, \hat{a}) = 0 \quad ; (39a)$$

$$A(X_i, Y_i, \hat{a}) = \tilde{V}_{11} - \tilde{U}_{20} \quad ; \quad (39b)$$

$$B(X_i, Y_i, \hat{a}) = \tilde{P}_{20} - \tilde{V}_{02} \quad ; \quad (39c)$$

$$C(X_i, Y_i, \hat{a}) = \tilde{U}_{02} - \tilde{P}_{11} \quad ; \quad (39d)$$

where  $Z_i$  stands for  $Z_1, Z_2, \dots, Z_n$ , and  $Z = X, Y$ .

- (2) Write the quadratic error propagation formula related to weighted and correlated measurements:

$$\left( \frac{\partial \phi}{\partial \hat{a}} \hat{\sigma}'_{\hat{a}} \right)^2 = \sum_{i=1}^n \left[ \left( \frac{\partial \phi}{\partial X_i} \right)^2 \frac{1}{w_{x_i}} + \left( \frac{\partial \phi}{\partial Y_i} \right)^2 \frac{1}{w_{y_i}} + \frac{2r_i}{\sqrt{w_{x_i}w_{y_i}}} \frac{\partial \phi}{\partial X_i} \frac{\partial \phi}{\partial Y_i} \right] . \quad (40)$$

- (3) Calculate the explicit expression of the partial derivatives,  $\partial \phi / \partial \hat{a}$ ,  $\partial \phi / \partial X_i$ ,  $\partial \phi / \partial Y_i$ ,  $1 \leq i \leq n$ , using Eqs. (31), (32), (35), (38), and (39).
- (4) Calculate the regression line slope variance estimator,  $(\hat{\sigma}_{\hat{a}})^2$ , using Eq. (40) and multiplying by the squared residual trace, using Eq. (11), and dividing by  $(n - 2)$ , as:

$$(\hat{\sigma}_{\hat{a}})^2 = \frac{(\hat{\sigma}'_{\hat{a}})^2}{n - 2} T_{\tilde{R}} \quad . \quad (41)$$

- (5) Cast the regression line intercept estimator, expressed by Eq. (26), into the implicit form:

$$\hat{b} = \psi(X_i, Y_i) = \tilde{Y} - \hat{a}(X_i, Y_i)\tilde{X} \quad ; \quad (42)$$

where  $Z_i$  stands for  $Z_1, Z_2, \dots, Z_n$ , and  $Z = X, Y$ .

- (6) Write the quadratic error propagation formula related to weighted and correlated measurements:

$$(\hat{\sigma}'_b)^2 = \sum_{i=1}^n \left[ \left( \frac{\partial \psi}{\partial X_i} \right)^2 \frac{1}{w_{x_i}} + \left( \frac{\partial \psi}{\partial Y_i} \right)^2 \frac{1}{w_{y_i}} + \frac{2r_i}{\sqrt{w_{x_i}w_{y_i}}} \frac{\partial \psi}{\partial X_i} \frac{\partial \psi}{\partial Y_i} \right]. \quad (43)$$

- (7) Calculate the explicit expression of the partial derivatives,  $\partial\psi/\partial X_i$ ,  $\partial\psi/\partial Y_i$ ,  $1 \leq i \leq n$ , using Eqs. (25), (42), and the theorem on the derivative of a function of a function:

$$\frac{\partial \phi}{\partial Z_i} = \frac{\partial \phi}{\partial \hat{a}} \frac{\partial \hat{a}}{\partial Z_i} \quad ; \quad Z = X, Y \quad ; \quad 1 \leq i \leq n \quad . \quad (44)$$

- (8) Calculate the regression line intercept variance estimator,  $(\hat{\sigma}_b)^2$ , using Eq. (43) and multiplying by the squared residual trace, using Eq. (11), and dividing by  $(n-2)$ , as:

$$(\hat{\sigma}_b)^2 = \frac{(\hat{\sigma}'_b)^2}{n-2} T_{\tilde{R}} \quad . \quad (45)$$

The calculation of the partial derivatives,  $\partial\phi/\partial\hat{a}$ ,  $\partial\phi/\partial X_i$ ,  $\partial\phi/\partial Y_i$ , and  $\partial\psi/\partial X_i$ ,  $\partial\psi/\partial Y_i$ ,  $1 \leq i \leq n$ , is performed in Appendix B.

Reasonable approximate values of the regression line slope and intercept variance estimators, are expressed as (Y66; Y69)<sup>3</sup>:

$$(\hat{\sigma}_{\hat{a}})^2 = \frac{1}{n-2} \frac{T_{\tilde{R}}}{\sum_{i=1}^n W_i (X_i - \tilde{X})^2} \quad ; \quad (46)$$

$$(\hat{\sigma}_b)^2 = (\hat{\sigma}_{\hat{a}})^2 \frac{\sum_{i=1}^n W_i (X_i)^2}{\sum_{i=1}^n W_i} \quad ; \quad (47)$$

---

<sup>3</sup> The numerator of the fraction in Eq. (47) has been omitted and put equal to unity in the parent paper (Y69) due to a printing error, as it can be argued by considering the physical dimensions on both sides, or by comparison with its counterpart in an earlier attempt (Y66).

which, using Eqs. (25), (35), and determining the explicit expression of the squared residual trace, may be cast under the equivalent form:

$$(\hat{\sigma}_{\hat{a}})^2 = \frac{1}{n-2} \left[ (\hat{a})^2 + \frac{\widetilde{W}_{02}}{\widetilde{W}_{20}} - 2\hat{a} \frac{\widetilde{W}_{11}}{\widetilde{W}_{20}} \right] ; \quad (48)$$

$$(\hat{\sigma}_{\hat{b}})^2 = (\hat{\sigma}_{\hat{a}})^2 (\widetilde{X^2}) ; \quad (49)$$

for a formal demonstration of Eq. (48) refer to Appendix C.

Relevant and useful special cases shall be discussed in the following subsections.

### 2.3 Uncorrelated errors in $X$ and in $Y$

In the limit of uncorrelated errors,  $(\sigma_{xy})_i = 0$ ,  $r_i = 0$ ,  $1 \leq i \leq n$ , Eqs. (22), (31), (32), and (38) reduce to:

$$W_i = \frac{w_{x_i} \Omega_i^2}{1 + a^2 \Omega_i^2} ; \quad 1 \leq i \leq n ; \quad (50)$$

$$V_i = \frac{w_{x_i} \Omega_i^4}{(1 + a^2 \Omega_i^2)^2} ; \quad 1 \leq i \leq n ; \quad (51)$$

$$U_i = 0 ; \quad 1 \leq i \leq n ; \quad (52)$$

$$P_i = \frac{w_{x_i} \Omega_i^2}{(1 + a^2 \Omega_i^2)^2} ; \quad 1 \leq i \leq n ; \quad (53)$$

accordingly, the pseudo cubic, Eq. (36), reduces to:

$$\widetilde{V}_{20} a^3 - 2\widetilde{V}_{11} a^2 - (\widetilde{W}_{20} - \widetilde{V}_{02}) a + \widetilde{W}_{11} = 0 ; \quad (54)$$

and the pseudo quadratic, Eq. (37), reduces to:

$$\widetilde{V}_{11} a^2 + (\widetilde{P}_{20} - \widetilde{V}_{02}) a - \widetilde{P}_{11} = 0 ; \quad (55)$$

where  $\widetilde{U}_{pq} = 0$  via Eqs. (35) and (52).

Following the same procedure outlined in the general case, the regression line slope and intercept estimators,  $\hat{a}$  and  $\hat{b}$ , and the regression line slope and intercept variance estimators,  $(\hat{\sigma}_{\hat{a}})^2$  and  $(\hat{\sigma}_{\hat{b}})^2$ , can be determined. For further details refer to the parent papers (Y66; Y69). A pictorial illustration of the method may be found in an additional paper (York, 1967). For a different but equivalent approach refer to an independent investigation (McIntyre et al., 1966). An earlier method implying approximate expressions is outlined in a pioneering attempt (Deming, 1943).

## 2.4 Completely correlated errors in $X$ and in $Y$

In the limit of completely correlated errors,  $(\sigma_{xy})_i = [(\sigma_{xx})_i(\sigma_{yy})_i]^{1/2}$  (where the positive and the negative root relate to correlation and anticorrelation, respectively),  $r_i = \text{sgn}(r_i)$ ,  $1 \leq i \leq n$ , Eqs. (19a) and (19b) reduce to:

$$X_i - x_i = [a\Omega_i - \text{sgn}(r_i)] \frac{\Omega_i \lambda_i}{w_{y_i}} ; \quad (56a)$$

$$Y_i - y_i = \text{sgn}(r_i) [a\Omega_i - \text{sgn}(r_i)] \frac{\lambda_i}{w_{y_i}} ; \quad (56b)$$

and the combination of Eqs. (56a) and (56b) yields:

$$\frac{Y_i - y_i}{X_i - x_i} = \frac{\text{sgn}(r_i)}{\Omega_i} ; \quad (57)$$

which is equivalent to Eq. (7) via Eq. (20). Accordingly, Eq. (11) reduces to:

$$T_{\tilde{R}} = \sum_{i=1}^n w_{x_i} (X_i - x_i)^2 = \sum_{i=1}^n w_{y_i} (Y_i - y_i)^2 ; \quad (58)$$

and the repetition of the procedure used in the general case yields again Eqs. (21)-(24), (26)-(33), (36)-(38), where  $r_i = \text{sgn}(r_i)$ ,  $1 \leq i \leq n$ , when necessary. In particular, Eqs. (22), (31), (32), and (38) reduce to:

$$W_i = \frac{w_{x_i} \Omega_i^2}{[a\Omega_i - \text{sgn}(r_i)]^2} ; \quad 1 \leq i \leq n ; \quad (59)$$

$$V_i = \frac{w_{x_i} \Omega_i^4}{[a\Omega_i - \text{sgn}(r_i)]^4} ; \quad 1 \leq i \leq n ; \quad (60)$$

$$U_i = \frac{w_{x_i} \Omega_i^3 \text{sgn}(r_i)}{[a\Omega_i - \text{sgn}(r_i)]^4} ; \quad 1 \leq i \leq n ; \quad (61)$$

$$P_i = \frac{w_{x_i} \Omega_i^2}{[a\Omega_i - \text{sgn}(r_i)]^4} ; \quad 1 \leq i \leq n ; \quad (62)$$

while the pseudo cubic and the pseudo quadratic maintain the formal expression of the general case, Eqs. (36) and (37), respectively. For a pictorial illustration of the method refer to the parent paper (Y69).

## 2.5 Errors in $X$ negligible with respect to errors in $Y$

In the limit of errors in  $X$  negligible with respect to errors in  $Y$ ,  $a^2(\sigma_{xx})_i \ll (\sigma_{yy})_i$ ,  $a(\sigma_{xy})_i \ll (\sigma_{yy})_i$ ,  $1 \leq i \leq n$ . Ideally,  $(\sigma_{xx})_i \rightarrow 0$ ,  $(\sigma_{xy})_i \rightarrow 0$ ,

$1 \leq i \leq n$ , which implies  $r_i \rightarrow 0$ ,  $w_{x_i} \rightarrow +\infty$ ,  $\Omega_i \rightarrow 0$ ,  $1 \leq i \leq n$ . Accordingly, the errors in  $X$  and in  $Y$  are uncorrelated.

In the limit,  $w_{x_i} \rightarrow +\infty$ ,  $1 \leq i \leq n$ , Eqs. (50)-(53), with due account taken of Eq. (20), reduce to (Y66):

$$W_i = w_{y_i} ; \quad 1 \leq i \leq n ; \quad (63)$$

$$V_i = 0 ; \quad 1 \leq i \leq n ; \quad (64)$$

$$U_i = 0 ; \quad 1 \leq i \leq n ; \quad (65)$$

$$P_i = w_{y_i} ; \quad 1 \leq i \leq n ; \quad (66)$$

accordingly, the pseudo cubic, Eq. (54), and the pseudo quadratic, Eq. (55), by use of Eq. (35) reduce to:

$$(\widetilde{w}_y)_{20}a - (\widetilde{w}_y)_{11} = 0 ; \quad (67)$$

and the regression line slope estimator reads:

$$\hat{a}_Y = \frac{(\widetilde{w}_y)_{11}}{(\widetilde{w}_y)_{20}} ; \quad (68)$$

finally, the substitution of Eq. (68) into (26) yields the regression line intercept estimator, as:

$$\hat{b}_Y = \widetilde{Y} - \hat{a}_Y \widetilde{X} ; \quad (69)$$

where the index, Y, stands for WLS(Y|X) i.e. weighted least square regression, or in particular OLS(Y|X) i.e. ordinary least square regression, of the dependent variable, Y, against the independent variable, X (Ial90).

With regard to the regression line slope and intercept variance estimators,  $(\hat{\sigma}'_{\hat{a}_Y})^2$  and  $(\hat{\sigma}'_{\hat{b}_Y})^2$ , with no account taken of the scatter of the data points,  $P_i \equiv (X_i, Y_i)$ , about the regression line, in the case under discussion Eqs. (40) and (43) reduce to:

$$\left( \frac{\partial \phi}{\partial \hat{a}_Y} \right)^2 (\hat{\sigma}'_{\hat{a}_Y})^2 = \sum_{i=1}^n \left( \frac{\partial \phi}{\partial Y_i} \right)^2 \frac{1}{w_{y_i}} ; \quad (70)$$

$$(\hat{\sigma}'_{\hat{b}_Y})^2 = \sum_{i=1}^n \left( \frac{\partial \psi}{\partial Y_i} \right)^2 \frac{1}{w_{y_i}} ; \quad (71)$$

and the substitution of Eqs. (208b), (209), and (210b) into (70) and (71), respectively, yields:

$$(\hat{\sigma}'_{\hat{a}_Y})^2 = \frac{1}{(\widetilde{w}_y)_{20}} ; \quad (72)$$

$$(\hat{\sigma}'_{\hat{b}_Y})^2 = (\hat{\sigma}'_{\hat{a}_Y})^2 \left[ \frac{(\widetilde{w}_y)_{20}}{(\widetilde{w}_y)_{00}} + (\widetilde{X})^2 \right] = (\hat{\sigma}'_{\hat{a}_Y})^2 (\widetilde{X}^2) ; \quad (73)$$



where Eqs. (25) and (35) have also been used.

The squared residual trace, expressed by Eq. (238), in the case under consideration via Eqs. (35) and (63) reduces to:

$$T_{\tilde{R}} = (\hat{a}_Y)^2 (\widetilde{w}_y)_{20} + (\widetilde{w}_y)_{02} - 2\hat{a}_Y (\widetilde{w}_y)_{11} \quad ; \quad (74)$$

and concerning the regression line slope and intercept variance estimators,  $(\hat{\sigma}_{\hat{a}_Y})^2$  and  $(\hat{\sigma}_{\hat{b}_Y})^2$ , with due account taken of the scatter of the data points,  $P_i \equiv (X_i, Y_i)$ , about the regression line, in the case under discussion Eqs. (41) and (45), by use of (63), (72), (73) and (74) reduce to:

$$(\hat{\sigma}_{\hat{a}_Y})^2 = \frac{1}{n-2} \left[ (\hat{a}_Y)^2 + \frac{(\widetilde{w}_y)_{02}}{(\widetilde{w}_y)_{20}} - 2\hat{a}_Y \frac{(\widetilde{w}_y)_{11}}{(\widetilde{w}_y)_{20}} \right] \quad ; \quad (75)$$

$$(\hat{\sigma}_{\hat{b}_Y})^2 = (\hat{\sigma}_{\hat{a}_Y})^2 \left[ \frac{(\widetilde{w}_y)_{20}}{(\widetilde{w}_y)_{00}} + (\bar{X})^2 \right] = (\hat{\sigma}_{\hat{a}_Y})^2 (\bar{X}^2) \quad ; \quad (76)$$

where Eqs. (25) and (35) have also been used. It can be seen that, in the case under consideration, the exact expression of the regression line slope and intercept estimators, Eqs. (75) and (76), coincide with the approximate expression in the general case, Eqs. (48) and (49), respectively.

The expression of the regression line slope and intercept estimators and related variance estimators, Eqs. (68), (69), (75), (76), coincide with their counterparts determined for WLS(Y|X) models in a recent attempt [Lavagnini and Magno, 2007, Eqs. (17)-(21) therein].

The substitution of Eq. (68) into (75) yields:

$$(\hat{\sigma}_{\hat{a}_Y})^2 = \frac{(\hat{a}_Y)^2}{n-2} \frac{D_{w_y}}{[(\widetilde{w}_y)_{11}]^2} \quad ; \quad (77)$$

$$D_{w_y} = (\widetilde{w}_y)_{02}(\widetilde{w}_y)_{20} - [(\widetilde{w}_y)_{11}]^2 \quad ; \quad (78)$$

where  $D_{w_y}$  is the determinant of a (weighted) deviation trace matrix, defined as a weighted deviation determinant.

Under the restriction of homoscedastic data,  $w_{x_i} = w_x$ ,  $w_{y_i} = w_y$ ,  $1 \leq i \leq n$ , which implies  $Q_i = Q$ ,  $Q = W, V, U, P$ , via Eqs. (22), (31), (32), (38), Eqs. (25) and (35) reduce to:

$$\tilde{Z} = \bar{Z} \quad ; \quad Z = X, Y \quad ; \quad (79a)$$

$$\sum_{i=1}^n (Z_i - \bar{Z}) = 0 \quad ; \quad Z = X, Y \quad ; \quad (79b)$$

$$\tilde{Q}_{pq} = Q S_{pq} \quad ; \quad Q = W, V, U, P \quad ; \quad (80)$$

$$S_{pq} = \sum_{i=1}^n (X_i - \bar{X})^p (Y_i - \bar{Y})^q \quad ; \quad (81)$$

where  $S_{pq}$  are the (unweighted) pure ( $p = 0$  and/or  $q = 0$ ) and mixed ( $p > 0$  and  $q > 0$ ) deviation traces, and  $S_{00} = n$ .

In the special case under discussion, Eqs. (68), (69), (75), (76), (77) and (78) reduce to:

$$\hat{a}_Y = \frac{S_{11}}{S_{20}} ; \quad (82)$$

$$\hat{b}_Y = \overline{Y} - \hat{a}_Y \overline{X} ; \quad (83)$$

$$(\hat{\sigma}_{\hat{a}_Y})^2 = \frac{1}{n-2} \left[ (\hat{a}_Y)^2 + \frac{S_{02}}{S_{20}} - 2\hat{a}_Y \frac{S_{11}}{S_{20}} \right] ; \quad (84)$$

$$(\hat{\sigma}_{\hat{b}_Y})^2 = (\hat{\sigma}_{\hat{a}_Y})^2 \left[ \frac{1}{n} S_{20} + (\overline{X})^2 \right] = (\hat{\sigma}_{\hat{a}_Y})^2 (\overline{X}^2) ; \quad (85)$$

$$(\hat{\sigma}_{\hat{a}_Y})^2 = \frac{(\hat{a}_Y)^2}{n-2} \frac{D_S}{(S_{11})^2} ; \quad (86)$$

$$D_S = S_{02} S_{20} - (S_{11})^2 ; \quad (87)$$

where  $D_S$  is the determinant of the (unweighted) deviation trace matrix, defined as the (unweighted) deviation determinant.

The expression of the regression line slope variance estimator, Eq. (84), coincides with earlier results known in literature in the special case of normal residuals [e.g., FB92, Eq. (4) therein, in the limit  $c^2 = (\sigma_{yy})_{zzz}/(\sigma_{xx})_{zzz} \rightarrow +\infty$ , where  $zzz = \text{ins, int}$ , denote instrumental and intrinsic scatter, respectively]. For a formal demonstration, see Appendix D. Then the expressions of the regression line slope and intercept variance estimators,  $(\hat{\sigma}_{\hat{a}_Y})^2$  and  $(\hat{\sigma}_{\hat{b}_Y})^2$ , reported above, hold provided the residuals obey a Gaussian distribution, as expected for functional models (Y66; Y69).

The expression of the regression line slope and intercept estimators and related variance estimators, Eqs. (82), (83), (84), (85), coincide with their counterparts determined for OLS( $Y|X$ ) models in a recent attempt [Lavagnini and Magno, 2007, Eqs. (3)-(7) therein].

## 2.6 Errors in $Y$ negligible with respect to errors in $X$

In the limit of errors in  $Y$  negligible with respect to errors in  $X$ ,  $(\sigma_{yy})_i \ll a^2(\sigma_{xx})_i$ ,  $(\sigma_{xy})_i \ll a(\sigma_{xx})_i$ ,  $1 \leq i \leq n$ . Ideally,  $(\sigma_{yy})_i \rightarrow 0$ ,  $(\sigma_{xy})_i \rightarrow 0$ ,  $1 \leq i \leq n$ , which implies  $r_i \rightarrow 0$ ,  $w_{y_i} \rightarrow +\infty$ ,  $\Omega_i \rightarrow +\infty$ ,  $1 \leq i \leq n$ . Accordingly, the errors in  $X$  and in  $Y$  are uncorrelated.

The model under discussion can be related to, but not confused with, the inverse regression, which has a large associate literature (e.g., Miller, 1966; Garden et al., 1980; Osborne, 1991; Brown, 1993; Lavagnini and Magno, 2007). More specifically, the inverse regression consists in the obtainement

of a variable,  $x$ , from an instrumental response,  $y$ , with the confidence interval for the true value of  $x$  (e.g., Brownlee, 1960; Lavagnini and Magno, 2007). A statistical calibration problem is a kind of inverse prediction, a problem of retrospection, and some authors call it inverse regression rather than calibration: it is probably best explained by considering a typical univariate calibration problem (Osborne, 1991).

In the limit,  $w_{y_i} \rightarrow +\infty$ ,  $1 \leq i \leq n$ , Eqs. (50)-(53), with due account taken of Eq. (20), reduce to (Y66):

$$W_i = a^{-2} w_{x_i} ; \quad 1 \leq i \leq n ; \quad (88)$$

$$V_i = a^{-4} w_{x_i} ; \quad 1 \leq i \leq n ; \quad (89)$$

$$U_i = 0 ; \quad 1 \leq i \leq n ; \quad (90)$$

$$P_i = 0 ; \quad 1 \leq i \leq n ; \quad (91)$$

accordingly, the pseudo cubic, Eq. (54), and the pseudo quadratic, Eq. (55), by use of Eq. (35) reduce to:

$$(\widetilde{w}_x)_{11} a - (\widetilde{w}_x)_{02} = 0 ; \quad (92)$$

and the regression line slope estimator reads:

$$\hat{a}_X = \frac{(\widetilde{w}_x)_{02}}{(\widetilde{w}_x)_{11}} ; \quad (93)$$

finally, the substitution of Eq. (93) into (26) yields the regression line intercept estimator, as:

$$\hat{b}_X = \widetilde{Y} - \hat{a}_X \widetilde{X} ; \quad (94)$$

where the index,  $X$ , stands for WLS( $X|Y$ ) i.e. weighted least square regression, or in particular OLS( $X|Y$ ) i.e. ordinary least square regression, of the independent variable,  $X$ , against the dependent variable,  $Y$  (Ial90).

With regard to the regression line slope and intercept variance estimators,  $(\hat{\sigma}'_{\hat{a}_X})^2$  and  $(\hat{\sigma}'_{\hat{b}_X})^2$ , with no account taken of the scatter of the data points,  $P_i \equiv (X_i, Y_i)$ , about the regression line, in the case under discussion Eqs. (40) and (43) reduce to:

$$\left( \frac{\partial \phi}{\partial \hat{a}_X} \right)^2 (\hat{\sigma}'_{\hat{a}_X})^2 = \sum_{i=1}^n \left( \frac{\partial \phi}{\partial X_i} \right)^2 \frac{1}{w_{x_i}} ; \quad (95)$$

$$(\hat{\sigma}'_{\hat{b}_X})^2 = \sum_{i=1}^n \left( \frac{\partial \psi}{\partial X_i} \right)^2 \frac{1}{w_{x_i}} ; \quad (96)$$

and the substitution of Eqs. (217a), (218), and (219a) into (95) and (96), respectively, yields:

$$(\hat{\sigma}'_{\hat{a}_X})^2 = \frac{(\hat{a}_X)^4}{(\widetilde{w}_x)_{02}} ; \quad (97)$$

$$(\hat{\sigma}'_{\hat{b}_X})^2 = \frac{(\hat{\sigma}'_{\hat{a}_X})^2}{(\hat{a}_X)^2} \left[ \frac{(\widetilde{w}_x)_{02}}{(\widetilde{w}_x)_{00}} + (\hat{a}_X)^2 (\widetilde{X})^2 \right] ; \quad (98)$$

where Eqs. (25) and (35) have also been used.

The squared residual trace, expressed by Eq. (238), in the case under consideration via Eqs. (35) and (88) reduces to:

$$T_{\widetilde{R}} = (\widetilde{w}_x)_{20} + (\hat{a}_X)^{-2} (\widetilde{w}_x)_{02} - 2(\hat{a}_X)^{-1} (\widetilde{w}_x)_{11} ; \quad (99)$$

and concerning the regression line slope and variance estimators,  $(\hat{\sigma}_{\hat{a}_X})^2$  and  $(\hat{\sigma}_{\hat{b}_X})^2$ , with due account taken of the scatter of the data points,  $\mathbf{P}_i \equiv (X_i, Y_i)$ , about the regression line, in the case under discussion Eqs. (41) and (45), by use of (88), (97), (98) and (99) reduce to:

$$(\hat{\sigma}_{\hat{a}_X})^2 = \frac{(\hat{a}_X)^2}{n-2} \left[ (\hat{a}_X)^2 \frac{(\widetilde{w}_x)_{20}}{(\widetilde{w}_x)_{02}} + 1 - 2\hat{a}_X \frac{(\widetilde{w}_x)_{11}}{(\widetilde{w}_x)_{02}} \right] ; \quad (100)$$

$$(\hat{\sigma}_{\hat{b}_X})^2 = \frac{(\hat{\sigma}_{\hat{a}_X})^2}{(\hat{a}_X)^2} \left[ \frac{(\widetilde{w}_x)_{02}}{(\widetilde{w}_x)_{00}} + (\hat{a}_X)^2 (\widetilde{X})^2 \right] ; \quad (101)$$

where it can be seen that, in the case under consideration, the exact expressions of the regression line slope and intercept estimator, Eqs. (100) and (101), respectively, do not coincide with the approximate expression in the general case, Eqs. (48) and (49), respectively.

The substitution of Eq. (93) into (100) yields:

$$(\hat{\sigma}_{\hat{a}_X})^2 = \frac{(\hat{a}_X)^2}{n-2} \frac{D_{w_x}}{[(\widetilde{w}_x)_{11}]^2} ; \quad (102)$$

$$D_{w_x} = (\widetilde{w}_x)_{02}(\widetilde{w}_x)_{20} - [(\widetilde{w}_x)_{11}]^2 ; \quad (103)$$

where  $D_{w_x}$  is a weighted deviation determinant.

In the special case of homoscedastic data,  $w_{x_i} = w_x$ ,  $1 \leq i \leq n$ , which implies  $Q_i = Q$ ,  $Q = W, V, U, P$ , via Eqs. (22), (31), (32), (38), and Eqs. (79) and (80) hold. Accordingly, Eqs. (93), (94), (100), (101), and (102) reduce to:

$$\hat{a}_X = \frac{S_{02}}{S_{11}} ; \quad (104)$$

$$\hat{b}_X = \bar{Y} - \hat{a}_X \bar{X} ; \quad (105)$$

$$(\hat{\sigma}_{\hat{a}_X})^2 = \frac{(\hat{a}_X)^2}{n-2} \left[ (\hat{a}_X)^2 \frac{S_{20}}{S_{02}} + 1 - 2\hat{a}_X \frac{S_{11}}{S_{02}} \right] ; \quad (106)$$

$$(\hat{\sigma}_{\hat{b}_X})^2 = \frac{(\hat{\sigma}_{\hat{a}_X})^2}{(\hat{a}_X)^2} \left[ \frac{1}{n} S_{02} + (\hat{a}_X)^2 (\bar{X})^2 \right] ; \quad (107)$$

$$(\hat{\sigma}_{\hat{a}_X})^2 = \frac{(\hat{a}_X)^2}{n-2} \frac{D_S}{(S_{11})^2} ; \quad (108)$$

where  $D_S$  is the deviation determinant, Eq. (87).

The expression of the regression line slope variance estimator, Eq. (106), is different from its counterpart calculated in an earlier attempt (FB92), due to the lack of an additional term which is negligible for  $D_S/(S_{11})^2 \ll 1$  and/or  $n \gg 1$ . For a formal demonstration, see Appendix D. Then the expressions of the regression line slope and intercept variance estimators,  $(\hat{\sigma}_{\hat{a}_X})^2$  and  $(\hat{\sigma}_{\hat{b}_X})^2$ , reported above, hold provided the residuals obey a Gaussian distribution, as expected for functional models (Y66; Y69), with the caveat due to the above mentioned discrepancy.

## 2.7 Generalized orthogonal regression

In the limit of constant  $y$  to  $x$  variance ratios and constant correlation coefficients, the following relations hold:

$$\frac{(\sigma_{yy})_i}{(\sigma_{xx})_i} = c^2 ; \quad \frac{w_{x_i}}{w_{y_i}} = \Omega_i^{-2} = c^2 ; \quad \frac{(\sigma_{xy})_i}{(\sigma_{xx})_i} = r_i c = r c ; \quad 1 \leq i \leq n ; \quad (109)$$

where the weights are assumed to be inversely proportional to related variances,  $w_{z_i} \propto 1/(\sigma_{zz})_i$ ,  $z = x, y$ , as usually done (e.g., FB92). It is worth noticing that Eq. (109) holds for both homoscedastic and heteroscedastic data. It can be seen that the lines of adjustment are oriented along the same direction (York, 1967) but are perpendicular to the regression line only in the special case,  $c^2 = 1$ , which is the genuine orthogonal regression (e.g., Carroll et al., 2006, Chap. 3, §3.4.2).

Earlier formulations of the model with respect to the parent paper (Y66) may be found in several attempts (e.g., Kummell, 1879; Koopmans, 1937; Deming, 1943; Tintner, 1945; Lindley, 1947; Anderson, 1951; Madansky, 1959) as well as later investigations (e.g., Barnett, 1967; Moran, 1971; Kendall and Stuart, 1979, Chap. 29; Fuller, 1980, 1987, Chap. 1, Sect. 1.3).

Taking into due account Eqs. (20) and (109), Eqs. (22), (31), (32) and (38) reduce to:

$$W_i = \frac{w_{x_i}}{a^2 + c^2 - 2rac} ; \quad 1 \leq i \leq n ; \quad (110)$$

$$V_i = \frac{w_{x_i}}{(a^2 + c^2 - 2rac)^2} ; \quad 1 \leq i \leq n ; \quad (111)$$

$$U_i = \frac{rcw_{x_i}}{(a^2 + c^2 - 2rac)^2} ; \quad 1 \leq i \leq n ; \quad (112)$$

$$P_i = \frac{c^2 w_{x_i}}{(a^2 + c^2 - 2rac)^2} ; \quad 1 \leq i \leq n ; \quad (113)$$

which, owing to Eq. (35), implies the following:

$$\tilde{Q}_{pq} = k_Q(\tilde{w}_x)_{pq} ; \quad Q = W, V, U, P ; \quad (114)$$

where  $k_Q = Q_i/w_{x_i}$  maintains constant.

Accordingly, the pseudo cubic, Eq. (36), and the pseudo quadratic, Eq. (37), reduce to:

$$[rc(\tilde{w}_x)_{20} - (\tilde{w}_x)_{11}]a^2 - [c^2(\tilde{w}_x)_{20} - (\tilde{w}_x)_{02}]a - [rc(\tilde{w}_x)_{02} - c^2(\tilde{w}_x)_{11}] = 0 \quad (115)$$

where  $r = \text{sgn}(r)$  in the limit of completely correlated errors in  $X$  and in  $Y$ .

In the special cases,  $c^2 \rightarrow +\infty$ ,  $c^2 \rightarrow 0$ , Eq. (115) reduces to (67) and (92), respectively, as expected. The regression line slope and intercept estimators,  $\hat{a}$  and  $\hat{b}$ , can be derived from Eqs. (115) and (26), respectively, where the parasite solution of the pseudo quadratic equation, in the former case, must be dismissed.

With regard to the regression line slope and intercept variance estimators,  $(\hat{\sigma}'_{\hat{a}_C})^2$  and  $(\hat{\sigma}'_{\hat{b}_C})^2$ , with no account taken of the scatter of the data points,  $P_i \equiv (X_i, Y_i)$ , about the regression line, in the case under discussion Eqs. (40) and (43) reduce to:

$$\left(\frac{\partial \phi}{\partial \hat{a}_C}\right)^2 (\hat{\sigma}'_{\hat{a}_C})^2 = \sum_{i=1}^n \left[ \left(\frac{\partial \phi}{\partial X_i}\right)^2 + c^2 \left(\frac{\partial \phi}{\partial Y_i}\right)^2 + 2rc \frac{\partial \phi}{\partial X_i} \frac{\partial \phi}{\partial Y_i} \right] \frac{1}{w_{x_i}} ; \quad (116)$$

$$(\hat{\sigma}'_{\hat{b}_C})^2 = \sum_{i=1}^n \left[ \left(\frac{\partial \psi}{\partial X_i}\right)^2 + c^2 \left(\frac{\partial \psi}{\partial Y_i}\right)^2 + 2rc \frac{\partial \psi}{\partial X_i} \frac{\partial \psi}{\partial Y_i} \right] \frac{1}{w_{x_i}} ; \quad (117)$$

and the substitution of Eqs. (226), (227), and (228), into (116) and (117), respectively, yields cumbersome expressions of the regression line slope and intercept variance estimators, which shall not be explicitly written here.

The squared residual trace, expressed by Eq. (238), in the case under consideration via Eqs. (35) and (110) reduces to:

$$T_{\tilde{R}} = \frac{(\hat{a}_C)^2(\tilde{w}_x)_{20} + (\tilde{w}_x)_{02} - 2\hat{a}_C(\tilde{w}_x)_{11}}{(\hat{a}_C)^2 + c^2 - 2rc\hat{a}_C} ; \quad (118)$$

and concerning the regression line slope and intercept variance estimators,  $(\hat{\sigma}_{\hat{a}_C})^2$  and  $(\hat{\sigma}_{\hat{b}_C})^2$ , with due account taken of the scatter of the data points,  $P_i \equiv (X_i, Y_i)$ , about the regression line, in the case under discussion Eqs. (41) and (45), by use of (118) may be cast into a more explicit form.

In the special case of uncorrelated errors,  $r \rightarrow 0$ , Eq. (115) reduces to:

$$(\widetilde{w}_x)_{11}a^2 + [c^2(\widetilde{w}_x)_{20} - (\widetilde{w}_x)_{02}]a - c^2(\widetilde{w}_x)_{11} = 0 \quad ; \quad (119)$$

which has the solutions (Deming, 1943):

$$\hat{a}_C = \frac{(\widetilde{w}_x)_{02} - c^2(\widetilde{w}_x)_{20}}{2(\widetilde{w}_x)_{11}} \left\{ 1 \mp \left[ 1 + c^2 \left( \frac{(\widetilde{w}_x)_{02} - c^2(\widetilde{w}_x)_{20}}{2(\widetilde{w}_x)_{11}} \right)^{-2} \right]^{1/2} \right\} \quad ; \quad (120)$$

and the regression line slope estimator is obtained disregarding the parasite solution. Then the substitution of Eq. (120) into (26) yields the regression line intercept estimator, as:

$$\hat{b}_C = \widetilde{Y} - \hat{a}_C \widetilde{X} \quad ; \quad (121)$$

where the index, C, denotes the case under discussion, with normal residuals (FB92).

The squared regression line slope estimator, via Eq. (120), reads:

$$(\hat{a}_C)^2 = \left\{ \frac{(\widetilde{w}_x)_{02} - c^2(\widetilde{w}_x)_{20}}{2(\widetilde{w}_x)_{11}} \mp c \left[ \frac{1}{c^2} \left( \frac{(\widetilde{w}_x)_{02} - c^2(\widetilde{w}_x)_{20}}{2(\widetilde{w}_x)_{11}} \right)^2 + 1 \right]^{1/2} \right\}^2 \quad ; \quad (122)$$

where the square root, if sufficiently close to unity, may be developed in binomial series and the terms of higher order neglected. The result is:

$$(\hat{a}_C)^2 = \left\{ \frac{(\widetilde{w}_x)_{02} - c^2(\widetilde{w}_x)_{20}}{2(\widetilde{w}_x)_{11}} \mp c \left[ 1 + \frac{1}{2} \frac{1}{c^2} \left( \frac{(\widetilde{w}_x)_{02} - c^2(\widetilde{w}_x)_{20}}{2(\widetilde{w}_x)_{11}} \right)^2 \right] \right\}^2 \quad ; \quad (123a)$$

$$\frac{1}{c^2} \left( \frac{(\widetilde{w}_x)_{02} - c^2(\widetilde{w}_x)_{20}}{2(\widetilde{w}_x)_{11}} \right)^2 \ll 1 \quad ; \quad (123b)$$

which, performing some algebra and neglecting the terms of higher order, takes the expression:

$$(\hat{a}_C)^2 = c^2 \mp 2c \frac{(\widetilde{w}_x)_{02} - c^2(\widetilde{w}_x)_{20}}{2(\widetilde{w}_x)_{11}} \quad ; \quad \frac{1}{c^2} \left( \frac{(\widetilde{w}_x)_{02} - c^2(\widetilde{w}_x)_{20}}{2(\widetilde{w}_x)_{11}} \right)^2 \ll 1 \quad ; \quad (124)$$

and additional algebra shows that the following relation holds:

$$\frac{(\hat{a}_C)^2 - (\widetilde{w}_x)_{02}/(\widetilde{w}_x)_{20}}{(\hat{a}_C)^2 - c^2} = 1 \pm \frac{1}{c} \frac{(\widetilde{w}_x)_{11}}{(\widetilde{w}_x)_{20}}; \quad \frac{1}{c^2} \left( \frac{(\widetilde{w}_x)_{02} - c^2(\widetilde{w}_x)_{20}}{2(\widetilde{w}_x)_{11}} \right)^2 \ll 1; \quad (125)$$

where further attention has to be devoted to the special case:

$$\frac{(\hat{a}_C)^2 - (\widetilde{w}_x)_{02}/(\widetilde{w}_x)_{20}}{(\hat{a}_C)^2 - c^2} = 1 \pm \lambda_{w_x} = 1 - \text{sgn}[(\widetilde{w}_x)_{11}] \lambda_{w_x} \quad ; \quad (126)$$

$$\lambda_{w_x} = \frac{(\widetilde{w}_x)_{11}}{[(\widetilde{w}_x)_{20}(\widetilde{w}_x)_{02}]^{1/2}} \quad ; \quad c^2 = \frac{(\widetilde{w}_x)_{02}}{(\widetilde{w}_x)_{20}} \quad ; \quad (127)$$

expressed in terms of the regression line correlation coefficient,  $\lambda_{w_x}$ .

With regard to the regression line slope and intercept variance estimators,  $(\hat{\sigma}'_{\hat{a}_C})^2$  and  $(\hat{\sigma}'_{\hat{b}_C})^2$ , the substitution of Eqs. (229), (233), and (234) into (116) and (117) particularized to uncorrelated errors,  $r = 0$ , yields after some algebra:

$$(\hat{\sigma}'_{\hat{a}_C})^2 = (\hat{a}_C)^2 \frac{(\widetilde{w}_x)_{02} + c^2(\widetilde{w}_x)_{20}}{[(\widetilde{w}_x)_{11}]^2} \quad ; \quad (128)$$

$$(\hat{\sigma}'_{\hat{b}_C})^2 = \frac{(\hat{a}_C)^2 + c^2}{(\widetilde{w}_x)_{00}} + (\hat{a}_C)^2 (\widetilde{X})^2 \frac{(\widetilde{w}_x)_{02} + c^2(\widetilde{w}_x)_{20}}{[(\widetilde{w}_x)_{11}]^2} \quad ; \quad (129)$$

where Eq. (35) has also been used.

The squared residual trace, expressed by Eq. (118), in the case under consideration reduces to:

$$T_{\widetilde{R}} = \frac{(\hat{a}_C)^2(\widetilde{w}_x)_{20} + (\widetilde{w}_x)_{02} - 2\hat{a}_C(\widetilde{w}_x)_{11}}{(\hat{a}_C)^2 + c^2} = \frac{(\hat{a}_C)^2(\widetilde{w}_x)_{20} - (\widetilde{w}_x)_{02}}{(\hat{a}_C)^2 - c^2} \quad ; \quad (130)$$

where Eq. (232) has also been used. Concerning the regression line slope and intercept variance estimators,  $(\hat{\sigma}_{\hat{a}_C})^2$  and  $(\hat{\sigma}_{\hat{b}_C})^2$ , Eqs. (41) and (45), by use of (128), (129) and (130), after a lot of algebra reduce to:

$$(\hat{\sigma}_{\hat{a}_C})^2 = \frac{1}{n-2} \frac{(\hat{a}_C)^2}{(\hat{a}_C)^2 - c^2} \frac{[(\widetilde{w}_x)_{02} + c^2(\widetilde{w}_x)_{20}][(\hat{a}_C)^2(\widetilde{w}_x)_{20} - (\widetilde{w}_x)_{02}]}{[(\widetilde{w}_x)_{11}]^2} \quad ; \quad (131)$$

$$(\hat{\sigma}_{\hat{b}_C})^2 = \frac{1}{n-2} \frac{1}{(\widetilde{w}_x)_{00}} \frac{(\hat{a}_C)^2 + c^2}{(\hat{a}_C)^2 - c^2} [(\hat{a}_C)^2(\widetilde{w}_x)_{20} - (\widetilde{w}_x)_{02}] + (\widetilde{X})^2 (\hat{\sigma}_{\hat{a}_C})^2 \quad ; \quad (132)$$

where further attention has to be devoted to the special case,  $c \rightarrow \hat{a}_C$ .

In the limit of errors in  $X$  negligible with respect to errors in  $Y$ ,  $c \rightarrow +\infty$ ,  $w_{x_i}/c^2 \rightarrow w_{y_i}$ ,  $\hat{a}_C \rightarrow \hat{a}_Y$ , it can be seen that Eqs. (131) and (132) reduce to (75) and (76), respectively, as expected.



In the limit of errors in  $Y$  negligible with respect to errors in  $X$ ,  $c \rightarrow 0$ ,  $c^2 w_{y_i} \rightarrow w_{x_i}$ ,  $\hat{a}_C \rightarrow \hat{a}_X$ , it can be seen that Eqs. (131) and (132) reduce to (100) and (101), respectively, as expected.

In the limit of the special case considered above,  $c^2 \rightarrow (\widetilde{w}_x)_{02}/(\widetilde{w}_x)_{20}$ , the combination of Eqs. (126) and (132) yields:

$$(\hat{\sigma}_{\hat{b}_C})^2 = \frac{1}{n-2} \frac{(\widetilde{w}_x)_{20}}{(\widetilde{w}_x)_{00}} [(\hat{a}_C)^2 + c^2] \{1 - \text{sgn}[(\widetilde{w}_x)_{11}] \lambda_{w_x}\} + (\widetilde{X})^2 (\hat{\sigma}_{\hat{a}_C})^2 ; \quad (133)$$

in terms of the regression line correlation coefficient,  $\lambda_{w_x}$ .

The parameter,  $c^2$ , appearing in Eqs. (131) and (132), may be eliminated via Eq. (232). The result is:

$$c^2 = \hat{a}_C \frac{(\widetilde{w}_x)_{02} - \hat{a}_C (\widetilde{w}_x)_{11}}{\hat{a}_C (\widetilde{w}_x)_{20} - (\widetilde{w}_x)_{11}} ; \quad (134)$$

$$(\hat{a}_C)^2 - c^2 = \frac{\hat{a}_C [(\hat{a}_C)^2 (\widetilde{w}_x)_{20} - (\widetilde{w}_x)_{02}]}{\hat{a}_C (\widetilde{w}_x)_{20} - (\widetilde{w}_x)_{11}} ; \quad (135)$$

and the substitution of Eqs. (134) and (135) into (131) yields after some algebra:

$$(\hat{\sigma}_{\hat{a}_C})^2 = \frac{(\hat{a}_C)^2}{n-2} \left\{ \frac{D_{w_x}}{[(\widetilde{w}_x)_{11}]^2} + \left[ \frac{D_{w_x}}{[(\widetilde{w}_x)_{11}]^2} + 2 - \frac{(\widetilde{w}_x)_{02} + (\hat{a}_C)^2 (\widetilde{w}_x)_{20}}{\hat{a}_C (\widetilde{w}_x)_{11}} \right] \right\} ; \quad (136)$$

where  $D_{w_x}$  is a weighted deviation determinant, Eq. (103).

In the special case of homoscedastic data,  $w_{x_i} = w_x$ ,  $1 \leq i \leq n$ , which implies  $Q_i = Q$ ,  $Q = W, V, U, P$ , via Eqs. (110)-(113), and Eqs. (79) and (80) hold. Accordingly, Eqs. (120), (121), (131), (132), (133) and (136) reduce to:

$$\hat{a}_C = \frac{S_{02} - c^2 S_{20}}{2S_{11}} \left\{ 1 \mp \left[ 1 + c^2 \left( \frac{S_{02} - c^2 S_{20}}{2S_{11}} \right)^{-2} \right]^{1/2} \right\} ; \quad (137)$$

$$\hat{b}_C = \overline{Y} - \hat{a}_C \overline{X} ; \quad (138)$$

$$(\hat{\sigma}_{\hat{a}_C})^2 = \frac{1}{n-2} \frac{(\hat{a}_C)^2}{(\hat{a}_C)^2 - c^2} \frac{[S_{02} + c^2 S_{20}][(\hat{a}_C)^2 S_{20} - S_{02}]}{(S_{11})^2} ; \quad (139)$$

$$(\hat{\sigma}_{\hat{b}_C})^2 = \frac{1}{n-2} \frac{1}{n} \frac{(\hat{a}_C)^2 + c^2}{(\hat{a}_C)^2 - c^2} [(\hat{a}_C)^2 S_{20} - S_{02}] + (\overline{X})^2 (\hat{\sigma}_{\hat{a}_C})^2 ; \quad (140)$$

$$(\hat{\sigma}_{\hat{b}_C})^2 = \frac{1}{n-2} \frac{1}{n} [(\hat{a}_C)^2 + c^2] [1 - \text{sgn}(S_{11}) \lambda_S] + (\overline{X})^2 (\hat{\sigma}_{\hat{a}_C})^2 ; \quad (141)$$

$$(\hat{\sigma}_{\hat{a}_C})^2 = \frac{(\hat{a}_C)^2}{n-2} \left[ \frac{D_S}{(S_{11})^2} + \left( \frac{D_S}{(S_{11})^2} + 2 - \frac{S_{02} + (\hat{a}_C)^2 S_{20}}{\hat{a}_C S_{11}} \right) \right] ; \quad (142)$$

where  $D_S$  is the deviation determinant, Eq. (87), and

$$\lambda_S = \frac{S_{11}}{[S_{20}S_{02}]^{1/2}} \quad ; \quad c^2 = \frac{S_{02}}{S_{20}} \quad ; \quad (143)$$

is the regression line correlation coefficient, and the above value of the parameter,  $c^2$ , relates to Eq. (141).

The expression of the regression line slope estimator, Eq. (137), coincides with its counterpart determined using the method of moments estimators [e.g., Fuller, 1987, Chap. 1, §1.3.2, Eq. (1.3.7) therein]. More specifically, the method of moments estimators and the least square estimators of  $a_C$  are the same (e.g., Fuller, 1987, Chap. 1, §1.3.3).

The expression of the regression line slope variance estimator, Eq. (142), is different from its counterpart calculated in an earlier attempt (FB92), due to a different term within round brackets on the right-hand side of Eq. (142). For a formal demonstration, see Appendix D. After long and strong work, it can be seen that the expression of both the regression line slope variance estimator determined in an earlier attempt (FB92) and intercept variance estimator expressed by Eq. (140), coincide with their counterparts reported in specific textbooks for structural models and uncorrelated errors (e.g., Fuller, 1987, Chap. 1, Sect. 1.3).

In the special case,  $c^2 = 1$ , the above results reduce to their counterparts related to genuine orthogonal regression, where the lines of adjustment are perpendicular to the regression line (Adcock, 1877, 1878; Pearson, 1901; Jones, 1937; Teissier, 1948; Kermack and Haldane, 1950).

Turning back to the general case of weighted residuals, but restricting to the special case,  $c^2 = (\widetilde{w}_x)_{02}/(\widetilde{w}_x)_{20}$ , the pseudo quadratic, Eq. (119), reduces to:

$$a^2 - c^2 = 0 \quad ; \quad (144)$$

which has the solutions (Kermack and Haldane, 1950; Y66):

$$\hat{a}_R = \mp \left[ \frac{(\widetilde{w}_x)_{02}}{(\widetilde{w}_x)_{20}} \right]^{1/2} \quad ; \quad (145)$$

and the regression line slope estimator is obtained disregarding the parasite solution. Then the substitution of Eq. (145) into (26) yields the regression line intercept estimator, as:

$$\hat{b}_R = \widetilde{Y} - \hat{a}_R \widetilde{X} \quad ; \quad (146)$$

where the index, R, denotes the case under discussion, with normal residuals (Y66).

The regression line slope and intercept variance estimators are obtained by substitution of Eq. (145) into (131), (133) and (136). After some algebra, the result is:

$$(\hat{\sigma}_{\hat{a}_R})^2 = \frac{2}{n-2} \frac{(\widetilde{w}_x)_{02}}{(\widetilde{w}_x)_{20}} ; \quad (147)$$

$$(\hat{\sigma}_{\hat{b}_R})^2 = \frac{2}{n-2} \frac{(\widetilde{w}_x)_{02}}{(\widetilde{w}_x)_{00}} \{1 - \text{sgn}[(\widetilde{w}_x)_{11}] \lambda_{w_x}\} + (\widetilde{X})^2 (\hat{\sigma}_{\hat{a}_R})^2 ; \quad (148)$$

$$(\hat{\sigma}_{\hat{a}_R})^2 = \frac{(\hat{a}_R)^2}{n-2} \left\{ \frac{D_{w_x}}{[(\widetilde{w}_x)_{11}]^2} + \left[ \frac{D_{w_x}}{[(\widetilde{w}_x)_{11}]^2} + 2 - 2\hat{a}_R \frac{(\widetilde{w}_x)_{20}}{(\widetilde{w}_x)_{11}} \right] \right\} ; \quad (149)$$

where  $D_{w_x}$  is a weighted deviation determinant, Eq. (103).

In the special case of homoscedastic data,  $w_{x_i} = w_x$ ,  $1 \leq i \leq n$ , which implies  $Q_i = Q$ ,  $Q = W, V, U, P$ , via Eqs. (110)-(113), and Eqs. (79) and (80) hold. Accordingly, Eqs. (145)-(149) reduce to:

$$\hat{a}_R = \mp \left( \frac{S_{02}}{S_{20}} \right)^{1/2} ; \quad (150)$$

$$\hat{b}_R = \overline{Y} - \hat{a}_R \overline{X} ; \quad (151)$$

$$(\hat{\sigma}_{\hat{a}_R})^2 = \frac{2}{n-2} \frac{S_{02}}{S_{20}} ; \quad (152)$$

$$(\hat{\sigma}_{\hat{b}_R})^2 = \frac{2}{n-2} \frac{S_{02}}{n} [1 - \text{sgn}(S_{11}) \lambda_S] + (\overline{X})^2 (\hat{\sigma}_{\hat{a}_R})^2 ; \quad (153)$$

$$(\hat{\sigma}_{\hat{a}_R})^2 = \frac{(\hat{a}_R)^2}{n-2} \left\{ \frac{D_S}{(S_{11})^2} + \left[ \frac{D_S}{(S_{11})^2} + 2 - 2\hat{a}_R \frac{S_{20}}{S_{11}} \right] \right\} ; \quad (154)$$

where  $D_S$  is the deviation determinant, Eq. (87).

The expression of the regression line slope variance estimator, Eq. (154), is different from its counterpart calculated in an earlier attempt (FB92), due to a different term within square brackets on the right-hand side of Eq. (154). For a formal demonstration, see Appendix D.

## 2.8 Extension to structural models

A nontrivial question is to what extent the above results, valid for functional models, can be extended to structural models. In general, assumptions related to structural models are different from their counterparts related to functional models (e.g., Buonaccorsi, 2006; 2010, Chap. 6, §6.4.5) but, on the other hand, they could coincide for a special subclass. In any case, whatever different assumptions and models can be made with regard to structural and functional models, results from the former are expected to tend to their

counterparts from the latter when the intrinsic scatter is negligible with respect to the instrumental scatter. It is worth noticing that most work on linear regression by astronomers involves the situation where both intrinsic scatter and heteroscedastic data are present (e.g., AB96; Tremaine et al., 2002; Kelly, 2007).

In structural models, a true point,  $P_i^*(x_i^*, y_i^*)$ , lying on the (true) regression line, is shifted by intrinsic scatter to an actual point,  $P_{Si}^*(x_{Si}, y_{Si})$ , outside the regression line. This last, in turn, is shifted by instrumental scatter to an observed point,  $P_i(X_i, Y_i)$ .

The coordinates of observed and actual points, according to Eq. (2), and the coordinates of actual and true points, are assumed to be related as:

$$(\xi_{F_z})_i = Z_i - z_{Si} \quad ; \quad Z = X, Y \quad ; \quad 1 \leq i \leq n \quad ; \quad (155)$$

$$(\xi_{S_z})_i = z_{Si} - z_i^* \quad ; \quad z = x, y; \quad 1 \leq i \leq n \quad ; \quad (156)$$

where the random variables,  $(\xi_{F_z})_i$ ,  $(\xi_{S_z})_i$ , obey the distributions,  $f_{F_z}[(\xi_{F_z})_i]$ ,  $f_{S_z}[(\xi_{S_z})_i]$ , respectively. While the former distribution is necessarily Gaussian, depending only on measurements processes, the latter distribution may be different, depending on a larger variety of processes.

With regard to true points, the substitution of Eq. (156) into (1) yields:

$$y_i^* + (\xi_{S_y})_i = a[x_i^* + (\xi_{S_x})_i] + b + \epsilon_i \quad ; \quad (157)$$

which, owing to Eq. (4), reduces to:

$$\epsilon_i = (\xi_{S_y})_i - a(\xi_{S_x})_i \quad ; \quad (158)$$

where the contribution of each variable to the intrinsic scatter is explicitly expressed. For the true regression line i.e. fixed slope,  $a$ , a null expectation value of  $\epsilon_i$  necessarily implies null expectation values of  $(\xi_{S_x})_i$  and  $(\xi_{S_y})_i$ ,  $1 \leq i \leq n$ , and vice versa.

With regard to observed points, the substitution of Eq. (155) into (1) yields:

$$Y_i - (\xi_{F_y})_i = a[X_i - (\xi_{F_x})_i] + b + \epsilon_i \quad ; \quad (159)$$

which, using Eq. (158), takes the form:

$$Y_i = aX_i + b + [(\xi_{S_y})_i + (\xi_{F_y})_i] - a[(\xi_{S_x})_i + (\xi_{F_x})_i] \quad ; \quad (160)$$

where the contribution of each variable to the intrinsic scatter is explicitly expressed.

A special subclass of structural models is defined according to the following assumptions.

- (a) The random variables,  $(\xi_{F_z})_i, (\xi_{S_z})_i, z = x, y, 1 \leq i \leq n$ , are independent.
- (b) The distributions,  $f_{S_z i}[(\xi_{S_z})_i], z = x, y, 1 \leq i \leq n$ , related to the intrinsic scatter, are Gaussian and corresponding expectation values are null.
- (c) The sum of variances related to the distributions,  $f_{F_z i}[(\xi_{F_z})_i], f_{S_z i}[(\xi_{S_z})_i], z = x, y, 1 \leq i \leq n$ , maintains constant as:

$$[(\sigma_{F_z})_i]^2 + [(\sigma_{S_z})_i]^2 = [(\sigma_z)_i]^2 = (\text{const}_z)_i; \quad z = x, y; \quad 1 \leq i \leq n. \quad (161)$$

Owing to assumption (b), the related distributions read:

$$f_{U_z i}[(\xi_{U_z})_i] = \frac{1}{\sqrt{2\pi}(\sigma_{U_z})_i} \exp \left\{ -\frac{[(\xi_{U_z})_i]^2}{2[(\sigma_{U_z})_i]^2} \right\} ;$$

$$U = F, S ; \quad z = x, y ; \quad 1 \leq i \leq n ; \quad (162)$$

with regard to the random variables,  $(\xi_{U_z})_i$ .

The random variables:

$$(\xi_z)_i = (\xi_{F_z})_i + (\xi_{S_z})_i = Z_i - z_i^*; \quad Z = X, Y; \quad z = x, y; \quad 1 \leq i \leq n; \quad (163)$$

via assumption (a) obey the distribution:

$$f_{z i}[(\xi_z)_i] = f_{F_z i}[(\xi_{F_z})_i] f_{S_z i}[(\xi_{S_z})_i] ; \quad (164)$$

which, using a theorem of statistics, via assumptions (a), (b), is also Gaussian, expressed as:

$$f_{z i}[(\xi_z)_i] = \frac{1}{\sqrt{2\pi}(\sigma_z)_i} \exp \left\{ -\frac{[(\xi_z)_i]^2}{2[(\sigma_z)_i]^2} \right\} ; \quad (165)$$

$$[(\sigma_z)_i]^2 = [(\sigma_{F_z})_i]^2 + [(\sigma_{S_z})_i]^2 ; \quad z = x, y ; \quad 1 \leq i \leq n ; \quad (166)$$

accordingly, residuals obey a Gaussian distribution and the weights,  $w_{x_i}, w_{y_i}, 1 \leq i \leq n$ , remain unchanged via assumption (c). Then, for a selected regression estimator, the regression line slope and intercept variance estimators are independent of the amount of instrumental and intrinsic scatter, including the limit of null intrinsic scatter (functional models) and null instrumental scatter (extreme structural models). In this view, the whole subclass of structural models under consideration could be related to functional modelling (Carroll et al., 2006, Chap. 2, §2.1).

## 3 An example of astronomical application

### 3.1 Astronomical introduction

With regard to stellar populations, the dependence of oxygen abundance on iron abundance, or  $[\text{O}/\text{H}]-[\text{Fe}/\text{H}]$  relation<sup>4</sup>, has been deeply investigated during the last decade (e.g., Carretta et al., 2000; Israelian et al., 2001a,b; Barbuy et al. (eds.), 2001; Jonsell et al., 2005; Fulbright et al., 2005; Garcia Perez et al., 2006; Melendez et al., 2006; Fabbian et al., 2009, hereafter quoted as Fal09; Rich and Boesgaard, 2009, hereafter quoted as RB09; Schmidt et al., 2009, hereafter quoted as Sal09).

It has been realized that the  $[\text{O}/\text{H}]-[\text{Fe}/\text{H}]$  relation is strongly dependent on both the selection of the spectroscopic oxygen lines and the choice of the atmosphere model. The discrepancy due to using different methods and different models remains large, and no general consensus on the best choice still exists. For further details refer to an earlier attempt (Caimmi, 2010).

Oxygen is the most abundant metal<sup>5</sup> in the universe, but it is more difficult than iron to detect. The population of available samples where oxygen abundances are directly determined, does not exceed a few hundreds at most (e.g., Ramirez et al., 2007; Melendez et al., 2008; RB09; Fal09; Sal09). Oxygen abundance in larger samples may be deduced by use of an inferred  $[\text{O}/\text{H}]-[\text{Fe}/\text{H}]$  relation.

According to the stellar evolution theory, oxygen is produced only via type II supernovae (SnII), characterized by massive ( $m \gtrsim 8m_{\odot}$ ) progenitors. On the contrary, iron is produced also via type Ia supernovae (SnIa), where a white dwarf ( $m < m_C$ ) attains the Chandrasekhar limit ( $m_C \approx 1.4m_{\odot}$ ) due to mass accretion from a close red giant companion, where white dwarfs are related to low-mass ( $m \lesssim 8m_{\odot}$ ) progenitors. During the lifetime of primeval SnII progenitors,  $\tau \lesssim 0.1\text{Gyr}$ , a linear  $[\text{O}/\text{H}]-[\text{Fe}/\text{H}]$  relation is expected, while the contrary holds at later times unless SnII contribution to iron production remains dominant. Three samples shall be used for bivariate least squares linear regression, namely RB09, Fal09, Sal09, where the denomination comes from related parent papers.

The RB09 sample ( $N = 49$ ) is made of a homogeneous subsample ( $N = 24$ ) of metal-poor ( $-3.5 < [\text{Fe}/\text{H}] < -2.2$ ) stars, and a non homogeneous sub-

---

<sup>4</sup>For a generic nuclide, N, the logarithmic number abundance is defined as  $[\text{N}/\text{H}] = \log(\text{N}/\text{H}) - \log(\text{N}/\text{H})_{\odot}$ , normalized to hydrogen, H, and to solar abundance,  $(\text{N}/\text{H})_{\odot}$ . The related mass abundance is defined as  $\phi_N = Z_N/(Z_N)_{\odot}$ . The relation:  $\log \phi_N = [\text{N}/\text{H}]$  holds to a good extent (Pagel, 1989; Malinie et al., 1993; Rocha-Pinto and Maciel, 1996; Caimmi, 2007).

<sup>5</sup>In astrophysical language, all elements heavier than helium are called “metals”.

sample ( $N = 25$ ) of higher-metallicity ( $-3.1 < [\text{Fe}/\text{H}] < -0.5$ ) stars. In both cases, the stellar population remains unspecified and oxygen abundance has been determined using standard local thermodynamical equilibrium (LTE) one-dimensional hydrostatic model atmospheres. Standard deviations are provided for each star, where typical values are  $\sigma_{[\text{Fe}/\text{H}]} = \sigma_{[\text{O}/\text{H}]} = 0.15$ . For further details refer to the parent paper (RB09).

The Fal09 sample ( $N = 44$ ) is made of halo stars ( $-3.3 < [\text{Fe}/\text{H}] < -1.0$ ) where oxygen abundance has been determined using three different methods involving (a) LTE one-dimensional hydrostatic model atmospheres; (b) three-dimensional hydrostatic model atmospheres in absence of LTE with no account taken of the inelastic collisions via neutral H atoms ( $S_{\text{H}} = 0$ ); (c) three-dimensional hydrostatic model atmospheres in absence of LTE with due account taken of the inelastic collisions via neutral H atoms ( $S_{\text{H}} = 1$ ). Standard deviations are not reported for each star, but typical values are mentioned to be  $\sigma_{[\text{Fe}/\text{H}]} = \sigma_{[\text{O}/\text{H}]} = 0.15$ . For further details refer to the parent paper (Fal09).

The RB09 and Fal09 samples have  $N = 11$  (necessarily halo) stars in common, where the values assumed for effective temperature and surface gravity have been determined using different methods, yielding different values for each star. For further details refer to an earlier attempt (Caimmi, 2010).

The Sal09 sample ( $N = 63$ ) is made of cool (late K and M) dwarfs ( $-1.8 < [\text{Fe}/\text{H}] < +0.2$ ) where oxygen abundance has been determined by use of the  $\gamma R_2$  0—0 TiO band at 7054 Å combined with previously derived abundances of Ti and Fe. Standard deviations are provided for each star, where typical values may be estimated as  $\sigma_{[\text{Fe}/\text{H}]} = \sigma_{[\text{O}/\text{H}]} = 0.15$ . A single star, LHS 185, is mentioned (Sal09, Table 1 therein) but excluded from further analysis. For further details refer to the parent paper (Sal09).

In conclusion, the Fal09 sample is made of homoscedastic data, while the remaining RB09 and Sal09 samples are made of heteroscedastic data. To a first extent, the latter may be considered as made of homoscedastic data where standard deviations are approximated to related typical values. Under the further assumption that intrinsic scatter is negligible with respect to instrumental scatter, i.e. functional models, the general results of section 2 may be particularized to the case under discussion, where errors in  $[\text{Fe}/\text{H}]$  and  $[\text{O}/\text{H}]$  may be considered as uncorrelated ( $r_i = 0$ ,  $1 \leq i \leq n$ ) to a good extent.

### 3.2 Statistical results

The  $[\text{O}/\text{H}]-[\text{Fe}/\text{H}]$  empirical relations are interpolated using the regression

Table 1: Regression line slope and intercept estimators,  $\hat{a}$  and  $\hat{b}$ , and related dispersion estimators,  $\hat{\sigma}_{\hat{a}}$ , and  $\hat{\sigma}_{\hat{b}}$ , for heteroscedastic models, G, Y, X, O, R, applied to the [O/H]-[Fe/H] empirical relation deduced from the following samples (from up to down): RB09, Sal09. Values related to different slope and intercept dispersion estimators (Y66) are also reported for comparison with current results (CRS). For G models, slope and intercept dispersion estimators were not evaluated in the present attempt. For Y models, different slope or intercept dispersion estimators yield coinciding values, as expected.

$m$	$\hat{a}$	$\hat{\sigma}_{\hat{a}}$		$\hat{b}$	$\hat{\sigma}_{\hat{b}}$		sample
		CRS	Y66		CRS	Y66	
G	0.7279		0.0294	+0.0043		0.0672	RB09
Y	0.6714	0.0314	0.0314	-0.1121	0.0675	0.0675	
X	0.7305	0.0290	0.0279	+0.0316	0.0735	0.0712	
O	0.6964	0.0278	0.0271	-0.0512	0.0707	0.0689	
R	0.7050	0.0282	0.0272	-0.0305	0.0725	0.0693	
G	0.6383		0.0435	+0.0619		0.0251	Sal09
Y	0.6167	0.0398	0.0398	+0.0439	0.0198	0.0198	
X	0.8652	0.0829	0.0664	+0.3080	0.0673	0.0575	
O	0.6355	0.0637	0.0541	+0.1461	0.0525	0.0469	
R	0.6927	0.0700	0.0560	+0.1864	0.0549	0.0485	

models, G, Y, X, O, R, for heteroscedastic data (RB09 and Sal09 samples) and Y, X, O, R, for homoscedastic data (Fal09 sample, cases LTE, SH0, SH1) and heteroscedastic data where instrumental scatters are taken equal to related typical values,  $\sigma_{[\text{Fe}/\text{H}]} = 0.15$ ,  $\sigma_{[\text{O}/\text{H}]} = 0.15$ , for both RB09 and Sal09 samples. Slope and intercept estimators together with related dispersion estimators are listed in Tables 1 and 2 for heteroscedastic and homoscedastic data, respectively. Also listed are values of slope and intercept dispersion estimators by earlier attempts (Y66; FB92) for comparison with their counterparts calculated in the current paper (CRS).

Owing to high difficulties intrinsic to the determination of slope and intercept dispersion estimators in the general case, related calculations were not performed in dealing with G models and only approximate expressions (Y66), Eqs. (48) and (49), were used. The regression line slope and intercept estimators are calculated using Eqs. (36) and (26), respectively. For



Table 2: Regression line slope and intercept estimators,  $\hat{a}$  and  $\hat{b}$ , and related dispersion estimators,  $\hat{\sigma}_{\hat{a}}$ , and  $\hat{\sigma}_{\hat{b}}$ , for homoscedastic models, Y, X, O, R, applied to the [O/H]-[Fe/H] empirical relation deduced from the following samples (from up to down): RB09, Sal09, Fal09, cases LTE, SH0, SH1. Values related to different slope and intercept dispersion estimators (Y66, FB92) are also reported for comparison with current results (CRS). For Y models, different slope or intercept dispersion estimators yield coinciding values, as expected.

$m$	$\hat{a}$	$\hat{\sigma}_{\hat{a}}$			$\hat{b}$	$\hat{\sigma}_{\hat{b}}$		sample
		CRS	FB92	Y66		CRS	Y66	
Y	0.6917	0.0317	0.0317	0.0317	-0.0766	0.0737	0.0737	FB09
X	0.7600	0.0348	0.0349	0.0332	+0.0742	0.0806	0.0773	
O	0.7143	0.0331	0.0327	0.0319	-0.0268	0.0766	0.0741	
R	0.7251	0.0336	0.0332	0.0321	-0.0030	0.0778	0.0746	
Y	0.5868	0.0461	0.0461	0.0461	+0.0908	0.0338	0.0338	Sal09
X	0.8077	0.0635	0.0637	0.0541	+0.2011	0.0430	0.0397	
O	0.6476	0.0526	0.0509	0.0468	+0.1212	0.0363	0.0343	
R	0.6885	0.0562	0.0541	0.0479	+0.1416	0.0381	0.0352	
Y	0.8961	0.0303	0.0303	0.0303	+0.5476	0.0663	0.0663	Fal09 (LTE)
X	0.9381	0.0317	0.0318	0.0310	+0.6366	0.0693	0.0678	
O	0.9150	0.0311	0.0310	0.0305	+0.5877	0.0680	0.0666	
R	0.9168	0.0312	0.0310	0.0305	+0.5916	0.0681	0.0667	
Y	1.2261	0.0432	0.0432	0.0432	+0.8717	0.0945	0.0945	Fal09 (SH0)
X	1.2884	0.0454	0.0454	0.0443	+1.0037	0.0991	0.0968	
O	1.2640	0.0448	0.0445	0.0436	+0.9519	0.0978	0.0953	
R	1.2569	0.0445	0.0443	0.0435	+0.9369	0.0973	0.0950	
Y	1.0492	0.0341	0.0341	0.0341	+0.6518	0.0745	0.0745	Fal09 (SH1)
X	1.0946	0.0356	0.0356	0.0348	+0.7479	0.0777	0.0761	
O	1.0732	0.0350	0.0349	0.0343	+0.7027	0.0765	0.0750	
R	1.0716	0.0350	0.0348	0.0343	+0.6993	0.0764	0.0750	

the remaining models, the regression line slope and intercept estimators and related dispersion estimators are calculated using Eqs. (68), (69), (75), (76), and (82), (83), (84), (85), case Y, heteroscedastic and homoscedastic data, respectively; Eqs. (93), (94), (100), (101), and (104), (105), (106), (107), case X, heteroscedastic and homoscedastic data, respectively; Eqs. (120), (121), (131), (132), and (137), (138), (139), (140), all related to the special value,  $c^2 = 1$  (genuine orthogonal regression), case O, heteroscedastic and homoscedastic data, respectively; Eqs. (145), (146), (147), (148), and (150), (151), (152), (153), case R, heteroscedastic and homoscedastic data, respectively.

The regression lines determined by use of the above mentioned methods are plotted in Figs. 1 and 2 for heteroscedastic and homoscedastic data, respectively, where sample denomination and population are indicated on each panel together with model captions. Homoscedastic data are conceived as a special case of heteroscedastic data in Fig. 1 to test the computer code, which is different for heteroscedastic and homoscedastic data. It can be seen that lower panels of Figs. 1 and 2 coincide, and the regression lines related to models G and O in lower panels of Figs. 1 and 2 also coincide, as expected. The whole set of regression lines for all methods and all samples is shown in the upper right panel of Figs. 1 and 2.

An inspection of Tables 1-2 and Figs. 1-2 discloses the following.

- (1) Either of the inequalities (Ial90):

$$\hat{a}_Y < \hat{a}_O < \hat{a}_R < 1 < \hat{a}_X ; \quad S_{11} > 0 ; \quad (167a)$$

$$\hat{a}_Y < 1 < \hat{a}_R < \hat{a}_O < \hat{a}_X ; \quad S_{11} > 0 ; \quad (167b)$$

holds for both heteroscedastic and homoscedastic data. In addition,  $\hat{a}_Y < \hat{a}_G < \hat{a}_X$  for heteroscedastic data, but a counterexample is provided in an earlier attempt (Y66).

- (2) Slope and intercept estimators by different methods are consistent within  $\mp\sigma$  for samples with lower dispersion (FB09, Fal09), while the contrary holds for samples with higher dispersion (Sal09), with regard to both heteroscedastic and homoscedastic data.
- (3) Slope and intercept dispersion estimators coincide with their counterparts related to earlier attempts for both heteroscedastic (Y66) and homoscedastic (Y66; FB92) data, in the special case of Y models. For the other models, the approximations exploited in an earlier attempt (Y66) make lower limits with respect to current results, while an alternative expression of the slope dispersion estimator (FB92) yields slightly different results.

- (4) Systematic variations due to different sample data are dominant with respect to the instrumental scatter.

In conclusion, regression lines deduced from different sample data represent correct (from the standpoint of regression models considered in the current attempt)  $[O/H]$ - $[Fe/H]$  relations, but no definitive choice can be made until systematic errors due to different methods and/or spectral lines in determining oxygen abundance, are alleviated.

## 4 Discussion

For an assigned sample, structural models belonging to a special subclass are indistinguishable from functional models provided restrictions (a)-(c) hold as outlined in subsection 2.8. Accordingly, the results of the current paper also apply to structural models of the kind considered. The expression of regression line slope and intercept estimators and related variance estimators in terms of weighted deviation traces, for heteroscedastic and homoscedastic data, makes a first step towards a unified formalism of bivariate least squares linear regression. The bisector method has not been dealt with in earlier attempts related to functional models (Y66; Y69), but only in later investigations involving extreme structural models (Ial90; FB92). For this reason, the bisector method has not been considered in the current paper.

Exact expressions of regression line slope and intercept estimators and related variance estimators have been determined from general formulae (Y69) in the limit of generalized orthogonal regression i.e.  $(\sigma_{yy})_i/(\sigma_{xx})_i = c^2$ ,  $1 \leq i \leq n$ . It is noteworthy that a constant variance ratio,  $c^2$ , for all data points, does not necessarily imply equal variances,  $(\sigma_{xx})_i = \sigma_{xx} = \text{const}$ ,  $(\sigma_{yy})_i = \sigma_{yy} = \text{const}$ ,  $1 \leq i \leq n$ . While regression line slope and intercept estimators attain a coinciding expression in different attempts (Y66; Y69; Ial90; FB92) with regard to a fixed model, the results of the current paper show that the contrary holds for related variance estimators.

Approximate expressions provided in earlier attempts (Y66; Y69) make (at least in computed cases) a lower limit to their exact counterparts, as shown in Tables 1-2. On the other hand, alternative expressions given in a later investigation, restricted to regression line slope variance estimators, yield different results (FB92).

The above mentioned discrepancy could be explained in different ways, namely: (1) calculations performed in the current paper were checked and repeated twice or more, but something wrong cannot be excluded; (2) the expression of the regression line slope variance estimator determined in the current paper, is approximate instead of exact, contrary to what reported in

the parent paper (Y69); (3) the expression of the regression line slope variance estimator used in an earlier attempt [FB92, Eq. (14) therein] for structural models with normal residuals and dominant intrinsic scatter, is approximate instead of exact; (4) for generalized orthogonal regression, the expression of the regression line slope variance estimator is different for functional (Y69) and structural (FB92) models even if, in the latter case, residuals obey a Gaussian distribution and the intrinsic scatter is dominant with respect to the instrumental scatter; (5) with regard to the regression line slope variance estimator, the method of partial differentiation, used in an earlier attempt (Y69) and in the current paper, yields different results with respect to the method of moments estimators [e.g., Fuller, 1987, Chap. 1, §1.3.2, Eq. (1.3.7) therein].

It is well known that the regression line slope and intercept estimators for Y models, Eqs. (82) and (83), are biased (e.g., Fuller, 1987, Chap. 1, §1.1.1; Carroll et al., 2006, Chap. 3, §3.2; Buonaccorsi, 2010, Chap. 4, §4.4). Biases can be explicitly expressed in the special case of homoscedastic models where random variables obey Gaussian distributions. More specifically, the condition  $1 - \rho_{20} \ll 1$  ensures bias effects are negligible, where  $\rho_{20}$  is the reliability ratio:

$$\rho_{20} = \frac{S_{20}}{S_{20} + (n - 1)\sigma_{xx}} ; \quad (168)$$

which implies  $0 \leq \rho_{20} \leq 1$ . For further details refer to specific monographies (e.g., Fuller, 1987, Chap. 1, §1.1.1; Carroll et al., 2006, Chap. 3, §3.2.1; Buonaccorsi, 2010, Chap. 4, §4.4). Accordingly, the reliability ratio cannot exceed unity, or in other words the regression line slope estimator, Eq. (82), is biased towards zero, as clearly shown in current literature (e.g., Kelly, 2007). Following a similar line of thought with regard to the regression line slope estimator for X models, Eq. (104), discloses the last is biased towards infinity, in the sense that the true slope is overestimated.

The regression line slope estimator for O models (genuine orthogonal regression) and R models (major-axis regression) lie between their counterparts related to Y and X models, according to Eq. (167), which implies bias corrections (e.g., Carroll et al., 2006, Chap. 3, §3.4.2). Though there is skepticism about an indiscriminate use of generalized orthogonal regression estimators, still it is accepted the method is viable provided both instrumental and intrinsic scatter are known (e.g., Carroll et al., 2006, Chap. 3, §3.4.2; Buonaccorsi, 2010, Chap. 4, §4.5).

With regard to heteroscedastic data, an inspection of Tables 1-2 shows that for lower data dispersion (RB09 sample) the values of regression line slope and intercept estimators, deduced for heteroscedastic (Table 1) and

homoscedastic (Table 2) data, are systematically smaller in the former case with respect to the latter, but remain consistent within  $\mp\sigma$ . For larger dispersion data (Sal09 sample) no systematic trend of the kind considered appears, but the values of regression line slope and intercept estimators are still consistent within  $\mp\sigma$  in different alternatives. It may be a general property of the regression models considered in the current attempt or, more realistically, intrinsic to the samples selected for the application performed in section 3.

The reliability ratio, Eq. (168), has been calculated for all sample data together with its counterpart for X models:

$$\rho_{02} = \frac{S_{02}}{S_{02} + (n - 1)\sigma_{yy}} \quad ; \quad (169)$$

and the inequalities,  $\rho_{20} > 0.92$ ,  $\rho_{02} > 0.91$ , hold in any case except  $\rho_{02} > 0.86$  for the Sal09 sample, which implies poorly biased regression line slope and intercept estimators for the samples considered using Y and X models and, a fortiori, using O and R models.

## 5 Conclusion

From the standpoint of a unified analytic formalism of bivariate least squares linear regression, functional models have been conceived as structural models where the intrinsic scatter is negligible (ideally null) with respect to the instrumental scatter.

Within the framework of classical error models, the dependent variable has been related to the independent variable according to the well known additive model (e.g., Carroll et al., 2006, Chap.1, §1.2, Chap.3, §3.2.1; Buonaccorsi, 2010, Chap.4, §4.3). Then the classical approach pursued in earlier papers (Y66; Y69) has been reviewed using a new formalism in terms of weighted deviation traces which, for homoscedastic data, reduce to usual quantities, leaving aside an unessential (but dimensional) multiplicative factor.

Regression line slope and intercept estimators, and related variance estimators, have been expressed in the general case of correlated errors in  $X$  and in  $Y$  for heteroscedastic data, and in the opposite limiting situations of (i) uncorrelated errors in  $X$  and in  $Y$ , and (ii) completely correlated errors in  $X$  and in  $Y$ . The special case of (C) generalized orthogonal regression has been considered in detail together with well known subcases, namely: (Y) errors in  $X$  negligible (ideally null) with respect to errors in  $Y$ ; (X) errors in  $Y$  negligible (ideally null) with respect to errors in  $X$ ; (O) genuine orthogonal regression; (R) reduced major-axis regression.

In the limit of homoscedastic data, the results determined for functional models have been compared with their counterparts related to extreme structural models i.e. the instrumental scatter is negligible (ideally null) with respect to the intrinsic scatter (Ial90; FB92). While regression line slope and intercept estimators for functional and structural models have necessarily been found to coincide, the contrary has been shown for related variance estimators even if the residuals obey a Gaussian distribution, with the exception of Y models.

An example of astronomical application has been considered, concerning the  $[O/H]$ - $[Fe/H]$  empirical relations deduced from five samples related to different stars and/or different methods of oxygen abundance determination. For selected samples and assigned methods, different regression models have been found to yield consistent results within the errors ( $\mp\sigma$ ) for both heteroscedastic and homoscedastic data. Conversely, it has been shown that samples related to different methods produce discrepant results, due to the presence of (still undetected) systematic errors, which implies no definitive statement can be made at present. A comparison has also been made between different expressions of regression line slope and intercept variance estimators, where fractional discrepancies were found to be not exceeding a few percent, which grows up to about 20% in presence of large dispersion data.

An extension of the results to structural models has been left to a forthcoming paper.

## 6 Note added in proof

The author is indebted with G.J. Babu and E.D. Feigelson for providing an earlier version of the erratum of their quoted paper (FB92) before publication (Feigelson and Babu, 2011). The quotation FB92 throughout the text has to be intended as including the original paper and the erratum (Feigelson and Babu, 1992, 2011).

## 7 Addendum

The slope variance estimator for generalized orthogonal regression with normal residuals, expressed in an earlier attempt [FB92, Eq.(4) therein, hereafter quoted as Eq.(FB4ev)] has been revised [FB92, erratum 2011, first equation therein, hereafter quoted as Eq.(FB4lv)]. Accordingly, the derivation of Eqs. (242), (251), (253), which started from Eq.(FB4ev), should be

repeated using Eq. (FB4lv). It can be seen the latter is closer to Eq. (142) than the former. The difference is small for the numerical application shown in Sect. 3.

## Acknowledgements

The author is grateful to an anonymous referee for enlightening comments which made substantial improvement to an earlier version of the current attempt. Thanks are due to G.J. Babu, E.D. Feigelson, M.A. Bershadsky, I. Lavagnini, S.J. Schmidt for fruitful e-mail correspondence on their quoted papers (FB92; AB96; Lavagnini and Magno, 2007; Sal09; respectively), and to C. Toniolo, C. Ghetti, S. Zoletto, V. Nascimbeni, for providing references on a consistent number of (non astronomical) statistical investigations quoted throughout the text.

## References

- [1] Adcock, R.J., 1877. *Analyst* 4, 183.
- [2] Adcock, R.J., 1878. *Analyst* 5, 53.
- [3] Akritas, M.G., Bershadsky, M.A., 1996. *ApJ* 470, 706 (AB96).
- [4] Anderson, T.W., 1951. *Ann. Math. Statist.* 22, 327.
- [5] Barbuy, B., Nissen, P.E., Peterson, R., Spite, F. (Ed.), *Proceedings of Oxygen abundances in Old Stars and Implications for Nucleosynthesis and Cosmology (IAU Joint Discussion 8)*, 2001. *New Astron. Rev.* 45, 509.
- [6] Barker, D.R., Diana, L.M., 1974. *Am. J. Phys.* 42, 224.
- [7] Barnett, V.D., 1967. *Biometrika* 54, 670.
- [8] Branham, R.L., 2001. *NewAR* 45, 649.
- [9] Brown, P.J., 1993. *Mesaurement, Regression and Calibration*, Oxford Statistical Science Series 12, Oxford Science Publications.
- [10] Brownlee, K.A., 1960. *Statistical Theory and Methodology in Science and Engineering*, New York, Wiley.
- [11] Buonaccorsi, J.P., 2006. *International Statistical Review* 74, 403.

- [12] Buonaccorsi, J.P., 2010. Masurement Error: Models, Methods and Applications, Chapman & Hall/CRC.
- [13] Caimmi, R., 2007. *NewA* 12, 289.
- [14] Caimmi, R., 2010. arXiv:1008.2057.
- [15] Carretta, E., Gratton, R.G., Sneden, C., 2000. *A&A* 356, 238.
- [16] Carroll, R.J., Roeder, K., Wasserman, L., 1999. *Biometrics* 55, 44.
- [17] Carroll, R.J., Ruppert, T.D., Stefanski, L.A., Crainiceanu, C.M., 2006. *Masurement Error in Nonlinear Models*, Monographs on Statistics and Applied Probability 105, ed. Chapman & Hall/CRC.
- [18] Clutton-Brock, M., 1967. *Technometrics* 9, 261.
- [19] Dellaportas, P., Stephens, D.A., 1995. *Biometrics* 51, 1085.
- [20] Deming, W.E., 1943. *Statistical Adjustment of Data*, J. Wiley & Sons, New York.
- [21] Fabbian, D., Nisseu, P.E., Asplund, M., et al., 2009. *A&A* 500, 1143 (Fal09).
- [22] Feigelson, E.D., Babu, G.J., 1992. *ApJ* 397, 55 (FB92).
- [23] Feigelson, E.D., Babu, G.J., 2011. *ApJ* 728, 72 (FB92, erratum).
- [24] Freedman, L.S., Fainberg, V., Kipnis, V., et al., 2004. *Biometrics* 60, 172.
- [25] Fulbright J.P., Rich R.M., McWilliam A., 2005. *NPA* 758, 197. Available from <astro-ph/0411041>.
- [26] Fuller, W.A., 1980. *Ann. Statist.* 8, 407.
- [27] Fuller, W.A., 1987. *Masurement Error Models*, ed. J. Wiley & Sons.
- [28] Garcia Perez, A.E., Asplund, M., Primas, F., et al., 2006. *A&A* 451, 621.
- [29] Garden, J.S., Mitchell, D.G., Mills, W.N., 1980. *Anal. Chem.* 52, 2310.
- [30] Gull, S.F., 1989. In *Maximum Entropy and Bayesian Methods*, ed. Skilling, Dordrecht, Kluwer, 511.



- [31] Isobe, T., Feigelson, E.D., Akritas, M.G., Babu, G.J., 1990. ApJ 364, 104 (Ial90).
- [32] Israelian, G., Rebolo, R., Garcia-Lopez, R.J., et al., 2001a. ApJ 551, 833.
- [33] Israelian, G., Rebolo, R., Garcia-Lopez, R.J., et al., 2001b. ApJ 560, 535.
- [34] Jones, H.E., 1937. Metron. 13, 21.
- [35] Jonsell, K., Edvardsson, B., Gustafsson, B., et al., 2005. A&A 440, 321.
- [36] Kelly, B.C., 2007. ApJ 665, 1489.
- [37] Kendall, M.G., Stuart, A., 1979. The Advanced Theory of Statistics, Vol. 2, 4th ed. Hafner, New York.
- [38] Kermack, K.A., Haldane, J.B.S., 1950. Biometrika 37, 30.
- [39] Klein, J.P., Moeschberger, M. L., 2005. Survival Analysis: Techniques for Censored and Truncated Data, 2nd edition, Springer.
- [40] Koopmans, T.C., 1937. Linear Regression Analysis of Economic Time Series, ed. DeErven F. Bohn, Haarlem, The Netherlands.
- [41] Kummell, C.H., 1879. Analyst 6, 97.
- [42] Lavagnini, I., Magno, F., 2007. Mass Spectrometry Rev. 26, 1.
- [43] Lindley, D.V., 1947. J. R. Statis. Soc. Suppl. 9, 218.
- [44] Madansky, A., 1959. J. Ann. Statis. Assoc. 54, 173.
- [45] Malinie, G., Hartmann, D.H., Clayton, D.D., Mathews, G.J., 1993. ApJ 413, 633.
- [46] McIntire, G.A., Brooks, C., Compston, W., Turek, A., 1966. J. Geophys. Res. 71, 5459.
- [47] Melendez, J., Shchukina, N.G., Vasiljeva, I.E., Ramirez, I., 2006. ApJ 642, 1082.
- [48] Melendez, J., Asplund, M., Alves-Brito, A., et al., 2008. A&A 484, L21.
- [49] Miller, R.P., 1966. Simultaneous Statistical Inference, New York, McGraw-Hill.

- [50] Moran, P.A.P., 1971. *J. Multivariate Anal.* 1, 232.
- [51] Osborne, C., 1991. *International Statistical Review* 59, 309.
- [52] Pagel, B.E.J., 1989. The G-dwarf Problem and Radio-active Cosmochronology. In: Beckman J.E., Pagel B.E.J. (eds.) *Evolutionary Phenomena in Galaxies*. Cambridge University Press, Cambridge (1989), p. 201.
- [53] Pearson, K., 1901. *Phil. Mag.* 2, 559.
- [54] Pourbaix, D., 1998. *A&ASS* 131, 377.
- [55] Press, W.M., Teukolsky, S.A., Wetterling, W.T., Flammery, B.P., 1992. *Numerical Recipes*, 2nd ed., Cambridge, Cambridge Univ. Press.
- [56] Ramirez, I., Allende Prieto, C., Lambert, D.L., 2007. *A&A* 465, 271.
- [57] Rich, J.A., Boesgaard, A.M., 2009. *ApJ* 701, 519 (RB09).
- [58] Rocha-Pinto, H.J., Maciel, W.J., 1996. *MNRAS* 279, 447.
- [59] Roy, S., Banerjee, T., 2006. *Ann. Inst. Statist. Math.* 58, 153.
- [60] Schafer, D.W., 1987. *Biometrika* 74, 385.
- [61] Schafer, D.W., 2001. *Biometrics* 57, 53.
- [62] Scheines, R., Hoijsink, H., Boosma, A., 1999. *Psychometrika* 64, 37.
- [63] Schmidt, S.J., Wallerstein, G., Woolf, V.M., Bean, J.L., 2009. *PASP* 121, 1083 (Sal09).
- [64] Schwarzenberg-Czarny, A., 1995. *A&ASS* 110, 405.
- [65] Teissier, G., 1948. *Biometrics* 4, 14.
- [66] Tintner, G., 1945. *Ann. Math. Statist.* 16, 304.
- [67] Tremaine, S., Ghehardt, K., Bender, R., et al., 2002. *ApJ* 574, 740.
- [68] York, D., 1966. *Canadian J. Phys.* 44, 1079 (Y66).
- [69] York, D., 1967. *Earth Plan. Science Lett.* 2, 479.
- [70] York, D., 1969. *Earth Plan. Science Lett.* 5, 320 (Y69).
- [71] Zellner, A., 1971. *An Introduction to Bayesian Inference in Econometrics*, New York, Wiley.

## Appendix

### A Solutions of cubic equations

Without loss of generality, a cubic equation may be cast under the standard form:

$$z^3 + a_1 z^2 + a_2 z + a_3 = 0 \quad ; \quad (170)$$

where the discriminant reads:

$$D = Q^3 + R^2 \quad ; \quad (171)$$

$$Q = \frac{3a_2 - a_1^2}{9} \quad ; \quad (172)$$

$$R = \frac{9a_1 a_2 - 27a_3 - 2a_1^3}{54} \quad ; \quad (173)$$

and one (the other two being complex conjugate) or three real solutions exist, according if  $D > 0$  or  $D \leq 0$ , respectively. The related expressions are:

$$z_1 = (R + D^{1/2})^{1/3} + (R - D^{1/2})^{1/3} - \frac{1}{3}a_1 \quad ; \quad (174)$$

for  $D > 0$  and:

$$z_1 = 2\sqrt{-Q} \cos\left(\frac{\theta}{3} + \frac{0\pi}{3}\right) - \frac{1}{3}a_1 \quad ; \quad (175)$$

$$z_2 = 2\sqrt{-Q} \cos\left(\frac{\theta}{3} + \frac{2\pi}{3}\right) - \frac{1}{3}a_1 \quad ; \quad (176)$$

$$z_3 = 2\sqrt{-Q} \cos\left(\frac{\theta}{3} + \frac{4\pi}{3}\right) - \frac{1}{3}a_1 \quad ; \quad (177)$$

$$\theta = \arctan \frac{\sqrt{-D}}{R} \quad ; \quad (178)$$

for  $D \leq 0$ , where  $\theta$  has to be replaced by  $\theta \mp \pi$  as appropriate, when a discontinuity of the arctan function occurs. A null factor has been put into Eq. (175) to save aesthetics.

### B Determination of partial derivatives

The explicit expressions of the regression line slope and intercept variance estimators, imply the calculation of the partial derivatives appearing in Eqs. (40) and (43), respectively. With regard to the regression line slope

variance estimator,  $(\hat{\sigma}'_a)^2$ , the slope estimator,  $\hat{a}$ , has to be considered as an independent variable in performing partial derivatives,  $\partial/\partial Z_i$ ,  $Z = X, Y$ . Using Eqs. (22), (25) and (35), the partial derivatives of the deviation from the weighted mean, with respect to the coordinates, read:

$$\frac{\partial}{\partial Z_i}(Z_\ell - \tilde{Z}) = \delta_{i\ell} - W'_i \quad ; \quad (179)$$

$$W'_i = \frac{W_i}{\bar{W}_{00}} \quad ; \quad \tilde{W}_{00} = \sum_{i=1}^n W_i = n\bar{W} \quad ; \quad (180)$$

where  $\delta_{i\ell}$  is the Kronecker symbol and the prime does not mean first derivation.

Using Eqs. (35) and (179), the partial derivatives of the weighted deviation traces, with respect to the coordinates, read:

$$\frac{\partial \tilde{Q}_{pq}}{\partial X_i} = pQ_i(X_i - \tilde{X})^{p-1}(Y_i - \tilde{Y})^q - pW'_i\tilde{Q}_{p-1,q} \quad ; \quad p > 0 \quad ; \quad (181a)$$

$$\frac{\partial \tilde{Q}_{pq}}{\partial Y_i} = qQ_i(X_i - \tilde{X})^p(Y_i - \tilde{Y})^{q-1} - qW'_i\tilde{Q}_{p,q-1} \quad ; \quad q > 0 \quad ; \quad (181b)$$

$$\frac{\partial \tilde{Q}_{0q}}{\partial X_i} = 0 \quad ; \quad \frac{\partial \tilde{Q}_{p0}}{\partial Y_i} = 0 \quad ; \quad (181c)$$

where the cases of interest are  $\tilde{Q}_{pq} = \tilde{V}_{11}, \tilde{U}_{20}, \tilde{P}_{20}, \tilde{V}_{02}, \tilde{P}_{11}, \tilde{U}_{02}$ .

Using Eqs. (22), (31), (32) and (38), the partial derivatives of the weights,  $W_i$ , with respect to the regression line slope estimator,  $\hat{a}$ , read:

$$\frac{\partial W_i}{\partial \hat{a}} = 2(U_i - \hat{a}V_i) \quad ; \quad (182)$$

and, in addition:

$$\frac{\partial V_i}{\partial \hat{a}} = 4(U'_i - \hat{a}V'_i) \quad ; \quad (183)$$

$$U'_i = \frac{W_i^3 r_i}{w_{x_i}^2 \Omega_i} \quad ; \quad (184)$$

$$V'_i = \frac{W_i^3}{w_{x_i}^2} \quad ; \quad (185)$$

$$\frac{\partial U_i}{\partial \hat{a}} = 4(U''_i - \hat{a}U'_i) \quad ; \quad (186)$$

$$U''_i = \frac{W_i^3 r_i^2}{w_{x_i}^2 \Omega_i^2} \quad ; \quad (187)$$

$$\frac{\partial P_i}{\partial \hat{a}} = 4(U_i''' - \hat{a}P_i') ; \quad (188)$$

$$U_i''' = \frac{W_i^3 r_i}{w_{x_i}^2 \Omega_i^3} ; \quad (189)$$

$$P_i' = \frac{W_i^3}{w_{x_i}^2 \Omega_i^2} ; \quad (190)$$

where the prime, the second, and the third, do not mean first, second, and third derivation, respectively.

Using Eqs. (35) and (183)-(190), the partial derivatives of related weighted deviation traces, with respect to the regression line slope estimator, read:

$$\frac{\partial \tilde{V}_{pq}}{\partial \hat{a}} = 4(\tilde{U}_{pq}' - \hat{a}\tilde{V}_{pq}') ; \quad (191)$$

$$\frac{\partial \tilde{U}_{pq}}{\partial \hat{a}} = 4(\tilde{U}_{pq}'' - \hat{a}\tilde{U}_{pq}') ; \quad (192)$$

$$\frac{\partial \tilde{P}_{pq}}{\partial \hat{a}} = 4(\tilde{U}_{pq}''' - \hat{a}\tilde{P}_{pq}') ; \quad (193)$$

which, together with Eq. (179), complete the set of results needed for an explicit expression of the partial derivatives of the function,  $\phi(X_i, Y_i, \hat{a})$ , defined by Eq. (39).

Using Eqs. (39) and (181), the explicit expression of the partial derivatives with respect to the coordinates, read:

$$\begin{aligned} \frac{\partial \phi}{\partial X_i} = & \left\{ [V_i(Y_i - \tilde{Y}) - W_i' \tilde{V}_{01}] - 2[U_i(X_i - \tilde{X}) - W_i' \tilde{U}_{10}] \right\} (\hat{a})^2 \\ & + 2[P_i(X_i - \tilde{X}) - W_i' \tilde{P}_{10}] \hat{a} - [P_i(Y_i - \tilde{Y}) - W_i' \tilde{P}_{01}] ; \end{aligned} \quad (194a)$$

$$\begin{aligned} \frac{\partial \phi}{\partial Y_i} = & [V_i(X_i - \tilde{X}) - W_i' \tilde{V}_{10}] (\hat{a})^2 - 2[V_i(Y_i - \tilde{Y}) - W_i' \tilde{V}_{01}] \hat{a} \\ & - \left\{ [P_i(X_i - \tilde{X}) - W_i' \tilde{P}_{10}] - 2[U_i(Y_i - \tilde{Y}) - W_i' \tilde{U}_{01}] \right\} ; \end{aligned} \quad (194b)$$

and using Eqs. (39), (191), (192), (193), the explicit expression of the partial derivative with respect to the regression line slope estimator reads:

$$\begin{aligned} \frac{\partial \phi}{\partial \hat{a}} = & 4(\tilde{U}_{20}' - \tilde{V}_{11}')(\hat{a})^3 + 4(\tilde{U}_{11}' - \tilde{U}_{20}'' - \tilde{P}_{20}' + \tilde{V}_{02}')(\hat{a})^2 \\ & + 2(2\tilde{U}_{20}''' - 2\tilde{U}_{02}' + 2\tilde{P}_{11}' - 2\tilde{U}_{02}'' + \tilde{V}_{11} - \tilde{U}_{20})\hat{a} \\ & + (4\tilde{U}_{02}'' - 4\tilde{U}_{11}''' + \tilde{P}_{20} - \tilde{V}_{02}) ; \end{aligned} \quad (195)$$

finally, the substitution of Eqs. (194) and (195) into (40) yields an explicit expression of the regression line slope variance estimator,  $(\hat{\sigma}'_a)^2$ , with no account taken of the scatter of the data points,  $\mathbf{P}_i \equiv (X_i, Y_i)$ , about the regression line (Y69).

With regard to the regression line intercept variance estimator,  $(\hat{\sigma}'_b)^2$ , the regression line slope estimator,  $\hat{a}$ , has to be considered as a function of the coordinates,  $\hat{a} = F_a(X_i, Y_i)$ ,  $1 \leq i \leq n$ , in performing partial derivatives,  $\partial/\partial Z_i$ ,  $Z = X, Y$ . Accordingly,  $\partial W_i/\partial Z_i \neq 0$ . Using Eq. (25), the partial derivatives of the weighted mean, with respect to the coordinates, after some algebra read:

$$\frac{\partial \tilde{Z}}{\partial Z_i} = \frac{1}{\tilde{W}_{00}} \left[ W_i + \sum_{\ell=1}^n \frac{\partial W_\ell}{\partial Z_i} (Z_\ell - \tilde{Z}) \right] ; \quad Z = X, Y ; \quad (196)$$

which, using Eqs. (35) and (182), together with a theorem on the partial derivatives of the function of a function:

$$\frac{\partial \phi}{\partial Z_i} = \frac{\partial \phi}{\partial \hat{a}} \frac{\partial \hat{a}}{\partial Z_i} ; \quad \frac{\partial W_\ell}{\partial Z_i} = \frac{\partial W_\ell}{\partial \hat{a}} \frac{\partial \hat{a}}{\partial Z_i} ; \quad 1 \leq i \leq n ; \quad (197)$$

after additional algebra takes the explicit form:

$$\frac{\partial \tilde{X}}{\partial X_i} = \frac{1}{\tilde{W}_{00}} \left[ W_i + 2 (\tilde{U}_{10} - \hat{a} \tilde{V}_{10}) \frac{\partial \phi / \partial X_i}{\partial \phi / \partial \hat{a}} \right] ; \quad (198a)$$

$$\frac{\partial \tilde{Y}}{\partial Y_i} = \frac{1}{\tilde{W}_{00}} \left[ W_i + 2 (\tilde{U}_{01} - \hat{a} \tilde{V}_{01}) \frac{\partial \phi / \partial Y_i}{\partial \phi / \partial \hat{a}} \right] ; \quad (198b)$$

and using Eqs. (42) and (198), the explicit expressions of the partial derivatives with respect to the coordinates read:

$$\begin{aligned} \frac{\partial \psi}{\partial X_i} &= -\hat{a} \frac{\partial \tilde{X}}{\partial X_i} - \tilde{X} \frac{\partial \hat{a}}{\partial X_i} \\ &= -\frac{\hat{a} W_i}{\tilde{W}_{00}} - \left[ \frac{2 \hat{a}}{\tilde{W}_{00}} (\tilde{U}_{10} - \hat{a} \tilde{V}_{10}) + \tilde{X} \right] \frac{\partial \phi / \partial X_i}{\partial \phi / \partial \hat{a}} ; \end{aligned} \quad (199a)$$

$$\begin{aligned} \frac{\partial \psi}{\partial Y_i} &= \frac{\partial \tilde{Y}}{\partial Y_i} - \tilde{X} \frac{\partial \hat{a}}{\partial Y_i} \\ &= \frac{W_i}{\tilde{W}_{00}} + \left[ \frac{2}{\tilde{W}_{00}} (\tilde{U}_{01} - \hat{a} \tilde{V}_{01}) - \tilde{X} \right] \frac{\partial \phi / \partial Y_i}{\partial \phi / \partial \hat{a}} ; \end{aligned} \quad (199b)$$

and the substitution of Eqs. (199) into (43) yields an explicit expression of the regression line intercept variance estimator,  $(\hat{\sigma}'_b)^2$ , with no account taken of the scatter of the data points,  $\mathbf{P}_i \equiv (X_i, Y_i)$ , about the regression line (Y69).

In the special case of errors in  $X$  negligible with respect to errors in  $Y$ , analysed in subsection 2.5, Eqs. (180), (184), (185), (187), (189), (190), via Eqs. (63)-(66) reduce to:

$$W'_i = \frac{w_{y_i}}{(\widetilde{w}_y)_{00}} \quad ; \quad (\widetilde{w}_y)_{00} = \sum_{i=1}^n w_{y_i} = n\overline{w}_y \quad ; \quad (200)$$

$$U'_i = 0 \quad ; \quad (201)$$

$$V'_i = 0 \quad ; \quad (202)$$

$$U''_i = 0 \quad ; \quad (203)$$

$$U'''_i = 0 \quad ; \quad (204)$$

$$P'_i = 0 \quad ; \quad (205)$$

and the particularization of Eq. (35) to the case under discussion via Eqs. (63)-(66), selecting  $Q = V, U, P$ ;  $p = 0, 1$ ;  $q = 1, 0$ ; and using Eq. (25), yields:

$$\widetilde{Q}_{10} = 0 \quad ; \quad (206)$$

$$\widetilde{Q}_{01} = 0 \quad ; \quad (207)$$

finally, the particularization of Eqs. (35), (194), (195), (199), to the case under discussion via Eqs. (63)-(66) and (200)-(207) produces:

$$\frac{\partial \phi}{\partial X_i} = w_{y_i} [2\hat{a}_Y(X_i - \widetilde{X}) - (Y_i - \widetilde{Y})] \quad ; \quad (208a)$$

$$\frac{\partial \phi}{\partial Y_i} = -w_{y_i}(X_i - \widetilde{X}) \quad ; \quad (208b)$$

$$\frac{\partial \phi}{\partial \hat{a}_Y} = (\widetilde{w}_y)_{20} \quad ; \quad (209)$$

$$\frac{\partial \psi}{\partial X_i} = -w_{y_i} \left\{ \frac{\hat{a}_Y}{(\widetilde{w}_y)_{00}} + \frac{\widetilde{X}}{(\widetilde{w}_y)_{20}} [2\hat{a}_Y(X_i - \widetilde{X}) - (Y_i - \widetilde{Y})] \right\} \quad ; \quad (210a)$$

$$\frac{\partial \psi}{\partial Y_i} = w_{y_i} \left[ \frac{1}{(\widetilde{w}_y)_{00}} + \frac{\widetilde{X}}{(\widetilde{w}_y)_{20}} (X_i - \widetilde{X}) \right] \quad ; \quad (210b)$$

and the substitution of Eqs. (208b), (209), and (210b) into (70) and (71) yields an explicit expression of the regression line slope and intercept variance estimator,  $(\hat{\sigma}'_{\hat{a}_Y})^2$  and  $(\hat{\sigma}'_{\hat{b}_Y})^2$ , respectively, with no account taken of the scatter of the data points,  $P_i \equiv (X_i, Y_i)$ , about the regression line, Eqs. (72) and (73).

In the special case of errors in  $Y$  negligible with respect to errors in  $X$ , analysed in subsection 2.6, Eqs. (180), (184), (185), (187), (189), (190), via Eqs. (88)-(91) reduce to:

$$W'_i = \frac{w_{x_i}}{(\widetilde{w}_x)_{00}} \quad ; \quad (\widetilde{w}_x)_{00} = \sum_{i=1}^n w_{x_i} = n\overline{w}_x \quad ; \quad (211)$$

$$U'_i = 0 \quad ; \quad (212)$$

$$V'_i = \frac{w_{x_i}}{(\hat{a}_X)^6} \quad ; \quad (213)$$

$$U''_i = 0 \quad ; \quad (214)$$

$$U'''_i = 0 \quad ; \quad (215)$$

$$P'_i = 0 \quad ; \quad (216)$$

and the particularization of Eq. (35) to the case under discussion via Eqs. (88)-(91), selecting  $Q = V, U, P$ ;  $p = 0, 1$ ;  $q = 1, 0$ ; and using Eq. (25), yields again Eqs. (206) and (207). Finally, the particularization of Eqs. (194), (195), (199), to the case under discussion via Eqs. (35), (88)-(91), (93), (206), (207), and (211)-(216) produces:

$$\frac{\partial \phi}{\partial X_i} = (\hat{a}_X)^{-2} w_{x_i} (Y_i - \widetilde{Y}) \quad ; \quad (217a)$$

$$\frac{\partial \phi}{\partial Y_i} = (\hat{a}_X)^{-3} w_{x_i} [\hat{a}_X (X_i - \widetilde{X}) - 2(Y_i - \widetilde{Y})] \quad ; \quad (217b)$$

$$\frac{\partial \phi}{\partial \hat{a}_X} = (\hat{a}_X)^{-4} (\widetilde{w}_x)_{02} \quad ; \quad (218)$$

$$\frac{\partial \psi}{\partial X_i} = -\hat{a}_X w_{x_i} \left[ \frac{1}{(\widetilde{w}_x)_{00}} + \frac{\hat{a}_X \widetilde{X}}{(\widetilde{w}_x)_{02}} (Y_i - \widetilde{Y}) \right] \quad ; \quad (219a)$$

$$\frac{\partial \psi}{\partial Y_i} = w_{x_i} \left\{ \frac{1}{(\widetilde{w}_x)_{00}} + \frac{\hat{a}_X \widetilde{X}}{(\widetilde{w}_x)_{02}} [2(Y_i - \widetilde{Y}) - \hat{a}_X (X_i - \widetilde{X})] \right\} \quad ; \quad (219b)$$

and the substitution of Eqs. (217a), (218), and (219a) into (95) and (96) yields an explicit expression of the regression line slope and intercept variance estimator,  $(\hat{\sigma}'_{\hat{a}_X})^2$  and  $(\hat{\sigma}'_{\hat{b}_X})^2$ , respectively, with no account taken of the scatter of the data points,  $\mathbf{P}_i \equiv (X_i, Y_i)$ , about the regression line, Eqs. (97) and (98).

In the special case of generalized orthogonal regression, analysed in subsection 2.7, Eqs. (180), (184), (185), (187), (189), (190), via Eqs. (110)-(113)



reduce to:

$$W'_i = \frac{w_{x_i}}{(\widetilde{w}_x)_{00}} \quad ; \quad (\widetilde{w}_x)_{00} = \sum_{i=1}^n w_{x_i} = n\overline{w}_x \quad ; \quad (220)$$

$$U'_i = \frac{rcw_{x_i}}{(a^2 + c^2 - 2rca)^3} \quad ; \quad (221)$$

$$V'_i = \frac{w_{x_i}}{(a^2 + c^2 - 2rca)^3} \quad ; \quad (222)$$

$$U''_i = \frac{r^2c^2w_{x_i}}{(a^2 + c^2 - 2rca)^3} \quad ; \quad (223)$$

$$U'''_i = \frac{rc^3w_{x_i}}{(a^2 + c^2 - 2rca)^3} \quad ; \quad (224)$$

$$P'_i = \frac{c^2w_{x_i}}{(a^2 + c^2 - 2rca)^3} \quad ; \quad (225)$$

and the particularization of Eq. (35) to the case under discussion via Eqs. (110)-(113), selecting  $Q = V, U, P$ ;  $p = 0, 1$ ;  $q = 1, 0$ ; and using Eq. (25), yields again Eqs. (206) and (207). Finally, the particularization of Eqs. (194), (195), (199), to the case under discussion via Eqs. (110)-(113), (206), (207), and (220)-(225) produces:

$$\frac{\partial\phi}{\partial X_i} = \frac{w_{x_i}[(\hat{a}_C)^2 - c^2](Y_i - \widetilde{Y}) + 2w_{x_i}\hat{a}_C(c - r\hat{a}_C)(X_i - \widetilde{X})}{[(\hat{a}_C)^2 + c^2 - 2rc\hat{a}_C]^2} \quad ; \quad (226a)$$

$$\frac{\partial\phi}{\partial Y_i} = \frac{w_{x_i}[(\hat{a}_C)^2 - c^2](X_i - \widetilde{X}) - 2w_{x_i}(\hat{a}_C - rc)(Y_i - \widetilde{Y})}{[(\hat{a}_C)^2 + c^2 - 2rc\hat{a}_C]^2} \quad ; \quad (226b)$$

$$\begin{aligned} \frac{\partial\phi}{\partial\hat{a}_C} &= [(\hat{a}_C)^2 + c^2 - 2rc\hat{a}_C]^{-3} \{ 2[rc(\widetilde{w}_x)_{20} - (\widetilde{w}_x)_{11}](\hat{a}_C)^3 \\ &\quad + [(\widetilde{w}_x)_{02} - c^2(\widetilde{w}_x)_{20}][3(\hat{a}_C)^2 - c^2] \\ &\quad + 2[c^2(\widetilde{w}_x)_{11} - rc(\widetilde{w}_x)_{02}][3\hat{a}_C - 2rc] \} \quad ; \quad (227) \end{aligned}$$

$$\begin{aligned} \frac{\partial\psi}{\partial X_i} &= -\frac{\hat{a}_C w_{x_i}}{(\widetilde{w}_x)_{00}} - \frac{\widetilde{X} w_{x_i}}{[(\hat{a}_C)^2 + c^2 - 2rc\hat{a}_C]^2} \left( \frac{\partial\phi}{\partial\hat{a}_C} \right)^{-1} \\ &\quad \times \{ [(\hat{a}_C)^2 - c^2](Y_i - \widetilde{Y}) - 2\hat{a}_C[r\hat{a}_C - c](X_i - \widetilde{X}) \} \quad ; \quad (228a) \end{aligned}$$

$$\begin{aligned} \frac{\partial\psi}{\partial Y_i} &= \frac{w_{x_i}}{(\widetilde{w}_x)_{00}} - \frac{\widetilde{X} w_{x_i}}{[(\hat{a}_C)^2 + c^2 - 2rc\hat{a}_C]^2} \left( \frac{\partial\phi}{\partial\hat{a}_C} \right)^{-1} \\ &\quad \times \{ [(\hat{a}_C)^2 - c^2](X_i - \widetilde{X}) - 2[\hat{a}_C - rc](Y_i - \widetilde{Y}) \} \quad ; \quad (228b) \end{aligned}$$

and the substitution of Eqs. (226), (227) and (228) into (116) and (117) yields an explicit expression of the regression line slope and intercept variance estimator,  $(\hat{\sigma}'_{\hat{a}_C})^2$  and  $(\hat{\sigma}'_{\hat{b}_C})^2$ , respectively, with no account taken of the scatter of the data points,  $\mathbf{P}_i \equiv (X_i, Y_i)$ , about the regression line.

In the limit of uncorrelated errors,  $r \rightarrow 0$ , Eqs. (226), (227) and (228) reduce to:

$$\frac{\partial \phi}{\partial X_i} = \frac{w_{x_i}}{[(\hat{a}_C)^2 + c^2]^2} \{[(\hat{a}_C)^2 - c^2](Y_i - \tilde{Y}) + 2\hat{a}_C c^2(X_i - \tilde{X})\} \quad ; \quad (229a)$$

$$\frac{\partial \phi}{\partial Y_i} = \frac{w_{x_i}}{[(\hat{a}_C)^2 + c^2]^2} \{[(\hat{a}_C)^2 - c^2](X_i - \tilde{X}) - 2\hat{a}_C(Y_i - \tilde{Y})\} \quad ; \quad (229b)$$

$$\frac{\partial \phi}{\partial \hat{a}_C} = \frac{[3(\hat{a}_C)^2 - c^2][(\tilde{w}_x)_{02} - c^2(\tilde{w}_x)_{20}] - 2\hat{a}_C[(\hat{a}_C)^2 - 3c^2](\tilde{w}_x)_{11}}{[(\hat{a}_C)^2 + c^2]^3} \quad ; \quad (230)$$

$$\begin{aligned} \frac{\partial \psi}{\partial X_i} = & -\frac{\hat{a}_C w_{x_i}}{(\tilde{w}_x)_{00}} - \tilde{X}[(\hat{a}_C)^2 + c^2] \\ & \times \frac{[(\hat{a}_C)^2 - c^2]w_{x_i}(Y_i - \tilde{Y}) + 2\hat{a}_C c^2 w_{x_i}(X_i - \tilde{X})}{[3(\hat{a}_C)^2 - c^2][(\tilde{w}_x)_{02} - c^2(\tilde{w}_x)_{20}] - 2\hat{a}_C[(\hat{a}_C)^2 - 3c^2](\tilde{w}_x)_{11}}; \end{aligned} \quad (231a)$$

$$\begin{aligned} \frac{\partial \psi}{\partial Y_i} = & \frac{w_{x_i}}{(\tilde{w}_x)_{00}} - \tilde{X}[(\hat{a}_C)^2 + c^2] \\ & \times \frac{[(\hat{a}_C)^2 - c^2]w_{x_i}(X_i - \tilde{X}) - 2\hat{a}_C w_{x_i}(Y_i - \tilde{Y})}{[3(\hat{a}_C)^2 - c^2][(\tilde{w}_x)_{02} - c^2(\tilde{w}_x)_{20}] - 2\hat{a}_C[(\hat{a}_C)^2 - 3c^2](\tilde{w}_x)_{11}}; \end{aligned} \quad (231b)$$

on the other hand, Eq. (119) may be cast under the equivalent form:

$$[(\hat{a}_C)^2 - c^2](\tilde{w}_x)_{11} = \hat{a}_C[(\tilde{w}_x)_{02} - c^2(\tilde{w}_x)_{20}] \quad ; \quad (232)$$

and the substitution of Eq. (232) into (230) and (231) yields:

$$\frac{\partial \phi}{\partial \hat{a}_C} = \frac{(\tilde{w}_x)_{11}}{\hat{a}_C[(\hat{a}_C)^2 + c^2]} \quad ; \quad (233)$$

$$\begin{aligned} \frac{\partial \psi}{\partial X_i} = & -\frac{\hat{a}_C w_{x_i}}{(\tilde{w}_x)_{00}} - \frac{\tilde{X} \hat{a}_C}{(\hat{a}_C)^2 + c^2} \\ & \times \frac{[(\hat{a}_C)^2 - c^2]w_{x_i}(Y_i - \tilde{Y}) + 2\hat{a}_C c^2 w_{x_i}(X_i - \tilde{X})}{(\tilde{w}_x)_{11}} \quad ; \end{aligned} \quad (234a)$$

$$\begin{aligned} \frac{\partial \psi}{\partial Y_i} = & \frac{w_{x_i}}{(\tilde{w}_x)_{00}} - \frac{\tilde{X} \hat{a}_C}{(\hat{a}_C)^2 + c^2} \\ & \times \frac{[(\hat{a}_C)^2 - c^2]w_{x_i}(X_i - \tilde{X}) - 2\hat{a}_C w_{x_i}(Y_i - \tilde{Y})}{(\tilde{w}_x)_{11}} \quad ; \end{aligned} \quad (234b)$$

finally, the substitution of Eqs. (229), (233) and (234) into (116) and (117) particularized to uncorrelated errors,  $r = 0$ , yields an explicit expression of the regression line slope and intercept variance estimator,  $(\hat{\sigma}'_{\hat{a}_C})^2$  and  $(\hat{\sigma}'_{\hat{b}_C})^2$ , respectively, with no account taken of the scatter of the data points,  $\mathbf{P}_i \equiv (X_i, Y_i)$ , about the regression line, Eqs. (128) and (129).

## C Regression line slope and intercept variance estimators

According to Eqs. (46) and (47), an explicit expression of the squared residual trace implies an explicit expression of the regression line slope and intercept variance estimators,  $(\hat{\sigma}_{\hat{a}})^2$  and  $(\hat{\sigma}_{\hat{b}})^2$ , respectively (Y69). The former may be obtained by substitution of Eqs. (19) and (21) into (11).

After some algebra, the result is:

$$T_{\tilde{R}} = \sum_{i=1}^n \frac{1}{1-r_i^2} \frac{W_i^2(\hat{a}X_i + \hat{b} - Y_i)^2}{w_{x_i}} \frac{1}{\Omega_i^2} \times \left[ (\hat{a}\Omega_i - r_i)^2 + (\hat{a}\Omega_i r_i - 1)^2 - 2r_i(\hat{a}\Omega_i - r_i)(\hat{a}\Omega_i r_i - 1) \right] \quad ; \quad (235)$$

where, in addition, the following relation holds:

$$\begin{aligned} & (\hat{a}\Omega_i - r_i)^2 + (\hat{a}\Omega_i r_i - 1)^2 - 2r_i(\hat{a}\Omega_i - r_i)(\hat{a}\Omega_i r_i - 1) \\ &= [(\hat{a})^2 \Omega_i^2 - 2\hat{a}\Omega_i r_i + 1](1 - r_i^2) = w_{x_i} \Omega_i^2 W_i^{-1} (1 - r_i^2) \quad ; \end{aligned} \quad (236)$$

where the former equality makes an identity and the latter is owing to Eq. (22).

The substitution of Eq. (236) into (235) yields:

$$T_{\tilde{R}} = \sum_{i=1}^n W_i(\hat{a}X_i + \hat{b} - Y_i)^2 \quad ; \quad (237)$$

and the combination of Eqs. (27), (35) and (237) produces:

$$T_{\tilde{R}} = (\hat{a})^2 \widetilde{W}_{20} + \widetilde{W}_{02} - 2\hat{a}\widetilde{W}_{11} \quad ; \quad (238)$$

which is the result of interest.

The substitution Eqs. (40), (194), (195) and (238) into (41) yields an explicit expression of the regression line slope variance estimator,  $(\hat{\sigma}_{\hat{a}})^2$ , with due account taken of the scatter of the data points,  $\mathbf{P}_i \equiv (X_i, Y_i)$ , about the regression line (Y69).

The substitution of Eqs. (43), (199), and (238) into (45) yields an explicit expression of the regression line intercept variance estimator,  $(\hat{\sigma}_b)^2$ , with due account taken of the scatter of the data points,  $\mathbf{P}_i \equiv (X_i, Y_i)$ , about the regression line (Y69).

If Eqs. (41) and (45) are approximated by Eqs. (46) and (47), respectively, the result is expressed by Eqs. (48) and (49), respectively.

In the special case of unweighted residuals,  $W_i = W$ ,  $\widetilde{W}_{pq} = WS_{pq}$ , Eq. (238) reduces to:

$$T_{\widetilde{R}} = W \left[ (\hat{a})^2 S_{20} + S_{02} - 2\hat{a}S_{11} \right] \quad ; \quad (239)$$

in terms of deviation traces.

## D Equivalence between earlier and current formulation

In the special case of homoscedastic functional or extreme structural models, where the ratio of instrumental or intrinsic dispersions in the two variables,  $Y$  and  $X$ , maintains constant,  $(\sigma_{yy})_i/(\sigma_{xx})_i = c^2$ ,  $1 \leq i \leq n$ , and the errors in  $Y$  and in  $X$  are uncorrelated,  $r_i = 0$ ,  $1 \leq i \leq n$ , the regression line slope estimator and the related dispersion estimator read:

$$\hat{a} = \frac{S_{02} - c^2 S_{20}}{2S_{11}} \left\{ 1 \mp \left[ 1 + c^2 \left( \frac{S_{02} - c^2 S_{20}}{2S_{11}} \right)^{-2} \right]^{1/2} \right\} \quad ; \quad (240)$$

$$(\hat{\sigma}_a)^2 = \frac{(\hat{a})^2}{(S_{11})^2} \left[ \left( \frac{S_{11}}{\hat{a}} + \frac{S_{02} - \hat{a}S_{11}}{c^2} \right) R - \frac{(\hat{a})^2}{n-1} \left( \frac{S_{02} - \hat{a}S_{11}}{c^2} \right)^2 \right] \quad ; \quad (241)$$

$$R = \frac{1}{n-2} \sum_{i=1}^n \left[ (Y_i - \overline{Y}) - \hat{a}(X_i - \overline{X}) \right]^2 \quad ; \quad (242)$$

under the further assumptions of unweighted normal residuals and large samples (FB92). With regard to Eq. (240), the plus instead of the double sign appears in the parent paper (FB92) but the latter is mentioned in an earlier paper (Ial90). In this case, the (physically meaningless) parasite solution must be disregarded.

The substitution of Eqs. (81) into (242) yields after some algebra:

$$R = \frac{1}{n-2} \left[ S_{02} + (\hat{a})^2 S_{20} - 2\hat{a}S_{11} \right] \quad ; \quad (243)$$

in terms of deviation traces.

The substitution of Eq. (239) into (243) yields:

$$R = \frac{1}{n-2} \frac{T_{\tilde{R}}}{W} ; \quad (244)$$

which shows the relation between different sums of squared residuals,  $R$  and  $T_{\tilde{R}}$ , via the dimensional constant,  $1/[(n-2)W]$ . A dimensionless counterpart to Eq. (244) reads:

$$\frac{R}{\hat{a}S_{11}} = \frac{1}{n-2} \frac{T_{\tilde{R}}}{\hat{a}\tilde{W}_{11}} ; \quad (245)$$

where  $\tilde{W}_{11} = WS_{11}$  in the case under consideration.

In the limit of errors in  $X$  negligible with respect to errors in  $Y$ ,  $c^2 \rightarrow +\infty$ , the square root on the right-hand side of Eq. (240) may be developed in binomial series with the terms of higher order neglected. The result is:

$$\hat{a}_Y = \frac{S_{11}}{S_{20}} ; \quad (246)$$

where the plus in the double sign on the right-hand side of Eq. (240) corresponds to a (physically meaningless) infinite value and for this reason has been disregarded, while the minus has been considered.

The combination of Eqs. (241) and (243), after neglecting the terms of higher order with respect to unity, yields:

$$(\hat{\sigma}_{\hat{a}_Y})^2 = \frac{1}{n-2} \frac{\hat{a}_Y}{S_{11}} \left[ S_{02} + (\hat{a}_Y)^2 S_{20} - 2\hat{a}_Y S_{11} \right] ; \quad (247)$$

and the substitution of Eq. (246) into (247) produces:

$$(\hat{\sigma}_{\hat{a}_Y})^2 = \frac{(\hat{a}_Y)^2}{n-2} \frac{D_S}{(S_{11})^2} ; \quad (248)$$

in terms of the deviation determinant, expressed by Eq. (87). It can be seen that Eqs. (246) and (248) coincide with (82) and (86), respectively, which implies the equivalence between earlier (FB92) and current (this paper) formulation, in the special case under discussion ( $c^2 \rightarrow +\infty$ ).

In the limit of errors in  $Y$  negligible with respect to errors in  $X$ ,  $c^2 \rightarrow 0$ , the term in square brackets on the right-hand side of Eq. (240) tends to unity. The result is:

$$\hat{a}_X = \frac{S_{02}}{S_{11}} ; \quad (249)$$

where the minus in the double sign on the right-hand side of Eq. (240) corresponds to a (physically meaningless) null value and for this reason has been disregarded, while the plus has been considered.

With this restriction, the square root on the right-hand side of Eq. (240) may be developed in binomial series with the terms of higher order neglected. After some algebra, the result is:

$$\frac{S_{02} - \hat{a}_X S_{11}}{c^2} = \frac{D_S}{S_{02}} ; \quad (250)$$

in terms of the deviation determinant, Eq. (87).

The substitution of Eq. (250) into (241) yields:

$$(\hat{\sigma}_{\hat{a}_X})^2 = \frac{(\hat{a}_X)^2}{(S_{11})^2} \left[ \frac{S_{11}R}{\hat{a}_X} + \frac{D_S R}{S_{02}} - \frac{(\hat{a}_X)^2 (D_S)^2}{n-1 (S_{02})^2} \right] ; \quad (251)$$

and the substitution of Eqs. (243) and (249) into (251) produces:

$$(\hat{\sigma}_{\hat{a}_X})^2 = \frac{(\hat{a}_X)^2}{n-2 (S_{11})^2} \left[ 1 + \frac{1}{n-1} \frac{D_S}{(S_{11})^2} \right] ; \quad (252)$$

in terms of the deviation determinant, expressed by Eq. (87). It can be seen that Eq. (249) coincides with Eq. (104), which implies the equivalence between earlier (FB92) and current (this paper) formulation, in the special case under discussion ( $c^2 \rightarrow 0$ ), concerning the regression line slope estimator,  $\hat{a}_X$ . The contrary holds for the regression line slope variance estimator,  $(\hat{\sigma}_{\hat{a}_X})^2$ , where Eq. (252) overestimates Eq. (108), the difference decreasing for increasing  $n$  and/or decreasing  $D_S/(S_{11})^2$ , and tends to be null in the limit,  $n \rightarrow +\infty$  and/or  $D_S/(S_{11})^2 \rightarrow 0$ .

Turning to the general case, it can be seen that Eq. (120) for unweighted residuals,  $w_{x_i} = w_x$ ,  $(\widetilde{w}_x)_{pq} = w_x S_{pq}$ , reduces to Eq. (240), which implies the validity of the pseudo quadratic, Eq. (119). The substitution of Eqs. (134) and Eq. (243) into (241) yields after a lot of algebra:

$$(\hat{\sigma}_{\hat{a}})^2 = \frac{(\hat{a})^2}{n-2} \left[ \frac{D_S}{(S_{11})^2} + \frac{1}{n-1} \left( \frac{\hat{a} S_{20}}{S_{11}} - 1 \right)^2 \right] ; \quad (253)$$

where in the limit,  $c \rightarrow +\infty$ ,  $c \rightarrow 0$ , Eq. (253) reduce to (248) and (252), respectively. A comparison between Eqs. (253) and (142) shows that earlier (FB92) and current (this paper) formulation are different due to the second term within square brackets.

In the special case,  $c^2 = S_{02}/S_{20}$ , which implies  $\hat{a} = (S_{02}/S_{20})^{1/2}$ , a similar conclusion is attained by comparison between Eqs. (253) and (154).

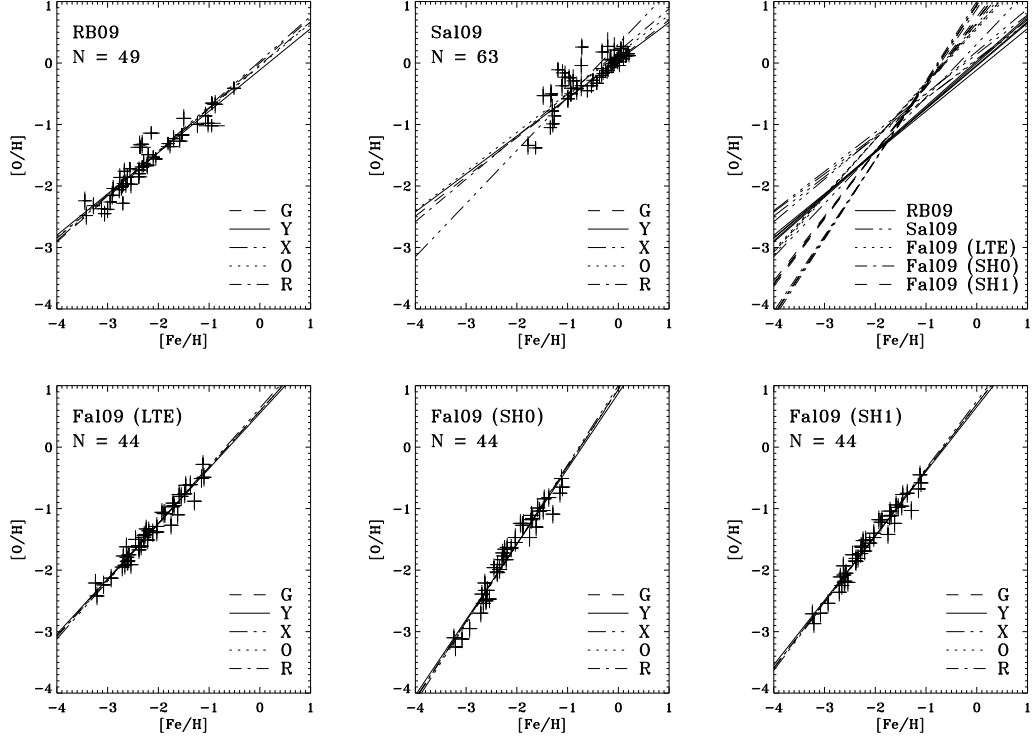


Figure 1: Regression lines related to  $[O/H]$ - $[Fe/H]$  empirical relations deduced from two samples with heteroscedastic data, RB09 and Sal09, and three samples with homoscedastic data (using the computer code for heteroscedastic data), Fa109, cases LTE, SH0, and SH1, indicated on each panel together with related population and model captions. The regression lines related to five different methods are shown for each sample on the top right panel. For further details refer to the text.

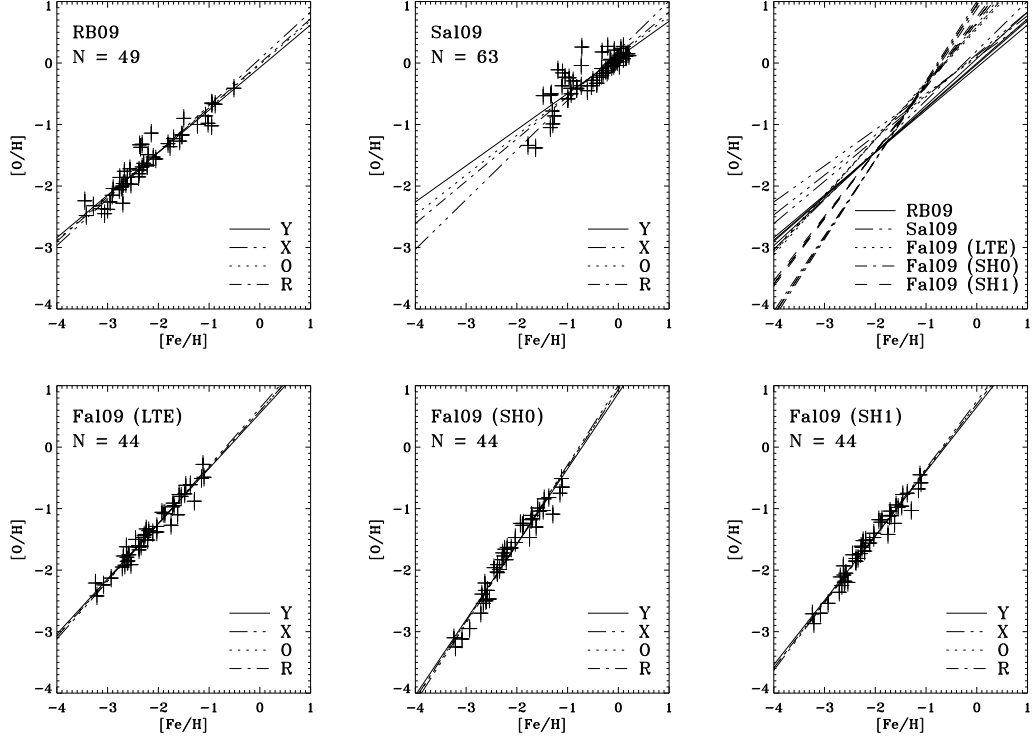


Figure 2: Regression lines related to  $[O/H]$ - $[Fe/H]$  empirical relations deduced from two samples with heteroscedastic data (with instrumental scatters taken equal to related typical values), RB09 and Sal09, and three samples with homoscedastic data, Fal09, cases LTE, SH0, and SH1, indicated on each panel together with related population and model captions. The regression lines related to four different methods are shown for each sample on the top right panel. For further details refer to the text.